

UNIVERSIDADE DE LISBOA  
FACULDADE DE CIÊNCIAS  
DEPARTAMENTO DE QUÍMICA E BIOQUÍMICA



# **BASE GENÉTICA DA HIPERCOLESTEROLEMIA FAMILIAR**

**Ana Catarina dos Santos Alves**

DOUTORAMENTO EM BIOQUÍMICA  
(ESPECIALIDADE EM GENÉTICA MOLECULAR)

2014



UNIVERSIDADE DE LISBOA  
FACULDADE DE CIÊNCIAS  
DEPARTAMENTO DE QUÍMICA E BIOQUÍMICA



# **BASE GENÉTICA DA HIPERCOLESTEROLEMIA FAMILIAR**

**Ana Catarina dos Santos Alves**

Tese orientada pela Doutora Mafalda Bourbon (Instituto Nacional de Saúde Dr. Ricardo Jorge) e pelo Professor Doutor Carlos Farinha (Faculdade de Ciências da Universidade de Lisboa), especialmente elaborada para a obtenção do grau de doutor em bioquímica, especialidade de Genética Molecular

2014



Ana Catarina Alves foi bolsreira de  
Doutoramento da Fundação para a Ciência e Tecnologia do  
Ministério da Ciência, Tecnologia e Ensino Superior  
SFRH/BD/27990/2006



## Prefácio

O trabalho de investigação descrito nesta tese de Doutoramento foi realizado na Unidade de Investigação e Desenvolvimento do Departamento de Promoção da Saúde e doenças Não transmissíveis do Instituto Nacional de Saúde Dr. Ricardo Jorge, sob a orientação da Doutora Mafalda Bourbon, investigadora auxiliar convidada neste Instituto, e co-orientação do Prof. Doutor Carlos Miguel Farinha, professor auxiliar na Faculdade de Ciências da Universidade de Lisboa. Os estudos funcionais foram desenvolvidos pela autora na Unidade de Biofísica (CSIC, UPV/EHU), Departamento de Bioquímica, Universidad del País Vasco, Bilbao, Spain.

As técnicas de pirosequenciação e sequenciação do exoma foram, realizados por empresas em regime de prestação de serviço às empresas: Biocante (pirosequenciação – *454 life science*), Progenika (pirosequenciação – *GS Junior System*) e Baseclear (sequenciação do exoma).

Este trabalho apresenta-se escrito em português de acordo com o novo acordo ortográfico.

De acordo com o disposto no artigo 40º do Regulamento de Estudos Pós-Graduados da Universidade de Lisboa, Deliberação nº961/2003, publicada no Diário da República – IIa Série, nº 153 de 5 de Julho de 2003, foram incluídos nesta tese resultados dos artigos abaixo indicados:

Bourbon M, Alves AC, Medeiros AM, Silva S, Soutar AK; on behalf of the investigators of the Portuguese FH study. Familial hypercholesterolaemia in Portugal. *Atherosclerosis*, 2008 Feb; 196(2):633-42.

Bourbon M, Duarte MA, Alves AC, Medeiros AM, Marques L, Soutar AK. 2009. Genetic diagnosis of familial hypercholesterolaemia: the importance of functional analysis of potential splice site mutations. *J Med Genet*. May;46(5):352-7.

Medeiros AM, Alves AC, Francisco V, Bourbon M. Update of the Portuguese Familial Hypercholesterolaemia Study. *Atherosclerosis*. 2010 Oct;212(2):553-8

**Alves AC**, Medeiros AM, Francisco V, Gaspar I, Rato Q, Bourbon M. Molecular diagnosis of Familial Hypercholesterolaemia - an important tool for cardiovascular risk stratification. *Rev Port Cardiol* 2010;29(06) 907-921

Silva S, **Alves AC**, Patel D, Malhó R, Soutar AK, Bourbon M. In vitro functional characterization of missense mutations in the LDLR gene. *Atherosclerosis*. 2012 Nov;225(1):128-34

**Alves AC**, Etxebarria A, Soutar AK, Martin C, Bourbon M. Novel functional APOB mutations outside LDL-binding region causing familial hypercholesterolaemia. *Hum Mol Genet*. 2014 Apr 1;23(7):1817-28. Epub 2013 Nov 13.

Medeiros AM, **Alves AC**, Aguiar P, Bourbon M. Cardiovascular risk assessment of dyslipidemic children: analysis of biomarkers for the correct assessment of monogenic dyslipidemia. *J Lipid Res*. 2014 Mar 13. Epub ahead of print

No cumprimento do disposto na referida deliberação, a autora assume a sua responsabilidade, exceto quando referido em contrário, na execução das experiências que permitiram a elaboração dos resultados apresentados, assim como na interpretação e discussão dos mesmos. Os resultados obtidos por outros foram incluídos com autorização dos mesmos para facilitar a compreensão dos trabalhos.

Outros artigos publicados em revistas nacionais contendo resultados obtidos durante o doutoramento:

MC Espinheira, C Vasconcelos, AM Medeiros, **AC Alves**, M Bourbon, A Guerra. Hypercholesterolemia – A disease with expression since childhood. *Rev Port Cardiol*. 2013; 32(5):379-386.

Artigos submetidos:

**Alves AC**, Medeiros AM and Bourbon M. Mutational analysis of a cohort with clinical diagnosis of familial hypercholesterolaemia: establishment of a data base for clinical practice (2014).



## Agradecimentos

Em primeiro lugar, gostaria de agradecer à Doutora Mafalda Bourbon, por me ter aceitado no seu laboratório e pela confiança depositada no meu trabalho. O seu entusiasmo e dedicação à investigação, em particular à Hipercolesterolemia Familiar, sempre foram uma fonte de inspiração para mim. As suas ideias e sugestões, bem como o interesse e entusiasmo, foram essenciais para a execução e desenvolvimento deste trabalho. Todo o conhecimento científico que me transmitiu permitiu-me crescer cientificamente, mas também como pessoa. Foi, e é, uma verdadeira mentora.

Ao Prof. Carlos Farinha, pelo apoio durante o doutoramento, principalmente na escrita da tese, na qual as suas sugestões e observações contribuíram para um melhor resultado.

Ao Prof. Cesar Martin, por me ter recebido no seu laboratório em Bilbao e proporcionado a oportunidade de realizar os estudos funcionais no gene *APOB* com a preciosa orientação do Aitor. A alegria, o entusiasmo e o empenho que vocês me transmitiram foram sem dúvida um momento alto deste trabalho.

À Doutora Astrid pela partilha científica, bem como a discussão de ideias e incentivo a novos desafios tecnológicos.

Agradeço à Fundação para a Ciência e a Tecnologia, por terem possibilitado a realização deste trabalho, através da concessão da bolsa de doutoramento de que fui recipiente entre 2008 e 2012, bem como através do projeto (FCT\_PTDC/SAU-GMG/101874/2008) que permitiu desenvolver grande parte deste trabalho de doutoramento.

A todas as minhas colegas do laboratório, que riram e desesperaram comigo, o meu obrigada. Em especial à Ana Medeiros, que esteve presente durante esta caminhada, obrigada pelas lágrimas que partilhaste comigo e pela compreensão em dias complicados. À Isabel Picanço, pelo seu entusiasmo em investigação, pela ajuda preciosa no envio das amostras para Bilbao e colheitas de sangue, bem como, nas sempre pertinentes questões que nos coloca. À Joana, Rita, Salima e Margarida, pela boa disposição nestes últimos meses. À enfermeira Ana Luísa Lourenço, pelas colheitas de sangue que fez no decurso deste trabalho.

A todos os meus amigos, um grande obrigada. Os jantares, os convívios, as vossas palavras amigas e conselhos ajudaram-me a recarregar as energias necessárias para o trabalho; sem vocês teria sido mais difícil. À minha «família», pelo apoio durante os jantares e nos jogos de cartas que ajudaram e muito a «respirar».

À Lurdes, pela revisão de algumas partes da tese, que tão importantes foram para mim. À Diana, pela ajuda preciosa nos últimos dias da tese, abdicando do seu sono para estar ao meu lado! À Sónia e ao João, obrigada por todas as gargalhadas que me proporcionam, mas também por todas as discussões científicas que tivemos.

À Rita, por estar sempre lá. Mesmo longe, conseguiste estar sempre presente na minha vida e dar todo o apoio emocional de que precisava. Sempre me fizeste acreditar em mim e nas minhas qualidades, mas nunca deixaste de me criticar quando achavas que era o momento certo, e isso fez-me crescer. Obrigada por todas as palavras sábias, mas principalmente pela tua amizade.

Aos irmãos Lapa, que acompanham sempre o meu percurso, dando-me ânimo, entusiasmo, mas principalmente a amizade necessária para nunca desistir dos meus objetivos, um grande obrigada.

Aos meus sobrinhos, em especial ao Rodrigo, pelo carinho e amor que me enchem a alma.

Por último, mas sempre os primeiros na minha vida, pai, mãe e mana, obrigada por tudo. Sem o vosso apoio, amor e compreensão, este trabalho não seria possível. Sempre acreditaram em mim, incentivaram e aturaram o mau humor (e que humor!) de dias sem inspiração ou de cansaço. Deram-me a força e confiança (às vezes apenas com um olhar) necessárias para nunca desistir, tornando possível esta caminhada. Aos meus pais, agradeço ainda o apoio financeiro que me deram ao longo da minha vida. Dedico-vos este trabalho.

## Índice

Prefácio .....	i
Agradecimentos .....	iii
Resumo .....	ix
Abstract .....	viii
ABREVIATURAS .....	xvi
I INTRODUÇÃO .....	1
I.1 Doenças genéticas .....	1
I.2 Hipercolesterolemia Familiar .....	2
I.2.1 Da observação clínica à descoberta do recetor das LDL .....	3
I.2.2 Via do Recetor das LDL .....	5
I.2.3 Outras causas de hipercolesterolemia .....	17
I.2.4 Caracterização molecular da FH .....	19
I.2.5 Tratamento .....	22
I.2.6 Estudo Português de Hipercolesterolemia Familiar .....	23
I.3 Mapeamento genético .....	23
I.3.1 Sequenciação de Sanger .....	24
I.3.2 <i>Next Generation Sequencing</i> .....	25
I.4 Objetivos .....	28
II MATERIAL E MÉTODOS .....	35
II.1 Recrutamento de doentes .....	35
II.2 Determinações bioquímicas .....	36
II.3 Colheita e processamento de amostras .....	36
II.3.1 Extração de DNA .....	37
II.3.2 Extração de RNA .....	38
II.4 Diagnóstico molecular da FH .....	39
II.4.1 Reorganização do diagnóstico molecular .....	39
II.4.2 Amplificação de DNA genómico por PCR .....	40

II.4.3	Análise dos fragmentos de DNA por eletroforese em gel de agarose .....	40
II.4.4	Sequenciação automática - método de Sanger.....	41
II.4.5	Análise de grandes rearranjos no gene <i>LDLR</i> por <i>multiplex ligation-dependent probe amplification</i> (MLPA) .....	42
II.5	<i>Next generation sequencing</i> .....	43
II.5.1	Pirosequenciação.....	43
II.5.2	Sequenciação de exoma .....	44
II.6	Programas de análise <i>in silico</i> .....	44
II.7	Estudos funcionais .....	45
II.7.1	Estudos de mutações <i>splicing</i> e promotor <i>LDLR</i> .....	45
II.7.2	Estudos de mutações missense no <i>LDLR</i> e no gene <i>APOB</i> .....	46
II.8	Painel de controlos normolipidémicos .....	51
II.9	Análise estatística .....	52
Parte 1 – Caracterização Molecular no âmbito do EPHF. ....		59
III.1.1	Período A .....	59
III.1.2	Período B .....	60
III.1.3	Pirosequenciação.....	65
III.1.2.2	<i>GS Junior System</i> .....	70
Parte 2 – Caracterização funcional de alterações encontradas nos genes associados a FH no âmbito do Estudo Português Hipercolesterolemia Familiar.....		77
III.2.1	Estudo funcional de alterações no promotor .....	77
III.2.2	Estudo funcional de alterações de <i>splicing</i> .....	79
III.2.3	Estudo funcional de alterações sinónimas .....	80
III.2.4	Estudo funcional de alterações pontuais no gene <i>LDLR</i> .....	84
III.2.5	Estudo funcional de ligação e internalização em alterações do gene <i>APOB</i> .....	87
III.2.6	Classificação das mutações .....	95
Parte 3 – Caracterização Bioquímica de doentes com diagnóstico clínico de FH –genótipo vs fenótipo .....		107
III.3.1	Grupo Pediátrico.....	108
III.3.2	Grupo Adultos .....	114

III.3.3 Caracterização clínica e bioquímica dos doentes com FH homozigótica e heterozigótica composta .....	121
Parte 4 – Sequenciação de exoma .....	127
IV.1 Caso índice 10146 .....	134
IV.2 Caso índice 25043 .....	136
IV.3 Caso índice 28120 .....	137
IV.4 Caso índice 10288 .....	138
IV DISCUSSÃO .....	145
IV.1 Caracterização molecular de casos índice com diagnóstico clínico de FH no âmbito do EPHF.....	145
IV.1.1 Pesquisa de mutações nos genes <i>LDLR</i> e <i>APOB</i> no âmbito do EPHF .....	146
IV.1.2 Pesquisa de mutações no gene <i>PCSK9</i> no âmbito do EPHF.....	147
IV.2 Pirosequenciação .....	150
IV.3 Estudos funcionais.....	154
IV.3.1 Estudos funcionais de alterações no gene <i>APOB</i> .....	154
IV.3.2 Estudos funcionais de alterações no gene <i>LDLR</i> .....	158
IV.4 Análise <i>in vitro</i> vs <i>in silico</i> .....	163
IV.5 Caracterização bioquímica - genótipo vs fenótipo .....	166
IV.5.1 Grupo Pediátrico .....	167
IV.5.2 Grupo Adultos .....	169
IV.5.3 FH homozigótica e heterozigótica composta .....	172
IV.6 Sequenciação de exoma .....	173
V CONCLUSÕES E PERSPETIVAS FUTURAS .....	185
REFERÊNCIAS.....	193



## Resumo

A Hipercolesterolemia Familiar (FH) é uma doença autossómica dominante caracterizada clinicamente por um aumento dos níveis de colesterol LDL no plasma, conduzindo à sua acumulação principalmente nos tendões (xantomas tendinosos) e nas artérias. Devido à acumulação de lípidos nas artérias, estes indivíduos desenvolvem aterosclerose muito cedo, tendo eventos cardiovasculares prematuramente. A FH apresenta uma frequência de 1/500, na forma heterozigótica, na maioria das populações europeias. Em Portugal estima-se que existam cerca de 20 000 casos de FH.

A identificação clínica da FH é possível, mas apenas o diagnóstico molecular da FH permite a correta identificação da doença, fundamentando a instituição de terapêutica farmacológica mais agressiva e/ou precoce, com a consequente redução do risco cardiovascular nos indivíduos afetados.

Geneticamente esta patologia caracteriza-se por mutações em três genes: *LDLR*, *APOB* e *PCSK9*. No entanto, a percentagem de indivíduos identificados molecularmente situa-se entre 30-80% dependendo da população estudada, uma vez que em Portugal apenas 40% dos doentes com diagnóstico clínico de FH apresentam uma mutação causadora de doença num dos três genes associados à FH.

O Estudo Português de Hipercolesterolemia Familiar (EPHF) tem como objetivos a realização de um estudo epidemiológico para a determinação da prevalência e distribuição da FH em Portugal, tendo implementado o estudo molecular desta doença. Para tal, a caracterização funcional de alterações encontradas em genes associados à FH é de extrema importância pois permite obter um diagnóstico definitivo de FH.

O objetivo deste trabalho doutoral era identificar e caracterizar a causa genética da hipercolesterolemia em famílias sem mutação identificável nos genes *LDLR* e *APOB* pelas metodologias usualmente utilizadas.

A primeira parte do presente foi dedicada à reorganização do diagnóstico molecular da FH e à pesquisa de mutações num novo gene – *PCSK9* –, à pesquisa de grandes rearranjos através da técnica de MLPA, bem como ao estudo dos genes *APOB* e *LDLRAP1* através de duas abordagens diferentes da técnica de pirosequenciação.

No total foram estudados no EPHF 642 casos índice (CI) (250 crianças e 392 adultos), tendo sido identificada uma alteração em 294 indivíduos. A reorganização do EPHF permitiu identificar 6 grandes rearranjos através da técnica de MLPA, identificando deste modo mais 19 indivíduos e 4 alterações no gene *PCSK9* em 5 CI. Através da técnica de pirosequenciação, identificaram-se 10 novas alterações potencialmente patogénicas no gene *APOB* em 10 indivíduos e 4 alterações no gene *LDLRAP1* em 4 CI.

A segunda parte do trabalho incluído nesta tese centrou-se na análise funcional de alterações cuja patogenicidade era desconhecida, uma vez que dos 294 CI com uma alteração apenas 192 CI apresentam uma mutação comprovadamente patogénica. Desta forma, foram analisadas funcionalmente para o gene *LDLR* 11 alterações de *splicing*, uma alteração no promotor e 4 alterações pontuais. Verificou-se que 8 alterações de *splicing* e 3 alterações *missense* eram patogénicas, e que 3 alterações de *splicing* e 1 alteração *missense* não originavam doença. No gene *APOB* foram estudadas funcionalmente 4 alterações, das quais apenas 2 destas são patogénicas. Estas mutações foram as primeiras descritas fora da região *consensus* de ligação do ligando ao LDLR postulada anteriormente.

Na impossibilidade da caracterização funcional de todas as alterações cujo estudo não tinha sido realizado, foi desenvolvida uma classificação com base em 6 programas bioinformáticos para as alterações *missense* e 3 programas para as alterações de *splicing*. Com base nesta classificação, identificaram-se 435 CI (175 CI e 260 familiares) com diagnóstico definitivo de FH, e 196 indivíduos (116 CI e 80 familiares) apresentavam um diagnóstico provável, 79 um diagnóstico possível e 46 neutro. Verificou-se que 594 indivíduos apresentavam uma alteração no gene *LDLR*, 30 indivíduos têm uma alteração no gene *APOB* e 7 indivíduos no gene *PCSK9*.

A terceira parte deste trabalho doutoral teve como objetivo correlacionar o genótipo vs com o fenótipo de todos os CI referenciados ao EPHF, tendo por base os diferentes diagnósticos moleculares que foram obtidos através da classificação das alterações com e sem estudos funcionais. A comparação dos fenótipos apresentados por crianças e adultos com diferentes tipos de mutação revelou que, embora em idade pediátrica o fenótipo não dependa do tipo de mutação, na idade adulta já existe uma clara diferença de fenótipos entre portadores de mutações nulas e *missense*. Esta observação reforça a importância do diagnóstico molecular precoce de modo a estratificar o risco cardiovascular no grupo pediátrico, uma vez que as crianças com mutações nulas



apresentam um maior risco cardiovascular, necessitando por isso de um aconselhamento especializado e implementação ainda mais precoce de terapêuticas hipolipemiantes.

A última parte deste trabalho focou-se na realização do estudo de sequenciação de exoma a 5 CI com diagnóstico clínico de FH, mas sem uma mutação identificada num dos 3 genes associados à FH. Uma vez que a sequenciação de exoma origina uma quantidade elevada de dados por amostra, de modo a restringir a análise, foi realizada uma seleção de 49 genes associados ao metabolismo lipídico, tendo-se verificado a existência de 20 alterações distribuídas por 17 genes. Através deste estudo, identificaram-se 2 alterações no gene *FLT1* em 2 CI não relacionados e 1 alteração no gene *SORT1* noutra CI que cossegregam com a hipercolesterolemia na família. Os estudos funcionais das alterações encontradas nestes genes serão importantes para classificar as mesmas quanto à sua patogenicidade. No caso de se comprovar que alguma destas alterações é funcional, esta será a primeira evidência de um quarto gene associado a FH.

Uma vez que a FH se encontra sub-diagnosticada em Portugal, o diagnóstico e aconselhamento genético da FH são importantes para a correta perceção e prevenção do risco familiar de DCV. Nas crianças e adolescentes, o diagnóstico genético é ainda mais importante, uma vez que se sabe que o risco cardiovascular é elevado, mas evitável, se medidas preventivas forem colocadas em prática. O futuro passa pela prevenção, em vez da resolução tardia das complicações cardiovasculares inerentes a esta patologia.

**Palavras-chaves:** Metabolismo lipídico, Hipercolesterolemia Familiar, estudos funcionais, patogenicidade, pirosequenciação, sequenciação do exoma

## Abstract

Familial Hypercholesterolemia (FH) is an autosomal dominant disorder characterized clinically by an increase in the levels of blood LDL-cholesterol that leads to lipid accumulation in tendons (xanthomas), premature atherosclerosis and coronary heart disease (CHD). FH has a heterozygous frequency of 1/500 in most European populations. It is estimated that Portugal should have about 20 000 cases of FH. The clinical identification of FH is possible but only the molecular diagnosis of FH allows the correct identification of the disease and justifies a more aggressive and/or early initiation of drug therapy, with consequent reduction in cardiovascular risk in affected individuals.

Genetically this disorder is characterized by mutations in three genes: *LDLR*, *APOB* and *PCSK9*. However the percentage of individuals identified molecularly is between 30-80% depending on the population studied; in Portugal only 40% of patients have an identified mutation, so other mutations in these genes or other gene defects should exist to explain the cause of hypercholesterolemia in the remaining families.

The general objective of the Portuguese Familial Hypercholesterolemia Study (PFHS) is to carry out an epidemiological study in order to determine the prevalence and distribution of FH in Portugal and to better understand the CHD. The main aim of the PFHS is to identify the genetic cause of the dyslipidemia in patients with clinical diagnosis of FH. The functional characterization of alterations found in genes associated with FH is extremely important for FH diagnosis because it leads to a definite diagnosis of FH.

The aim of this doctoral work was to identify and characterize the genetic cause of hypercholesterolemia in families without identified mutations in *LDLR* and *APOB* genes by methods commonly used.

The first part of this doctoral work consisted in the reorganization of the molecular diagnosis of FH, the detection of mutations in a novel gene, the *PCSK9* gene, to the screening of large rearrangements by MLPA, and the sequencing of *APOB* and *LDLRAP1* using two different approaches other than pyrosequencing.

The PFHS identified a total of 642 patients (250 children and 392 adults) and an alteration was identified in 294 individuals. The reorganization of PFHS allowed the identification of 6 major rearrangements by MLPA technique, thereby identifying 19

more individuals with FH; 4 alterations in *PCSK9* gene in 5 index patients. Through pyrosequencing technique 10 alterations in *APOB* gene in 10 individuals and 4 alterations in gene *LDLRAP1* in 4 patients were identified.

In a second part of the project, the work was focused on the functional analysis of alterations with unknown pathogenicity, as among 294 patients with an alteration only 192 patients have a proven pathogenic mutation. So, functional studies were performed for 11 splicing alterations in the *LDLR* gene, one alteration in the promoter region and 4 missense alterations. It was found that 8 splicing alterations and 3 missense alterations were pathogenic and 3 splicing alterations and one missense were not cause of disease. Functional studies were performed for 4 alterations found in the *APOB* gene, 2 of which were identified as pathogenic. These mutations were the first mutations ever described outside the previously postulated consensus region of the ligand binding to the LDLR.

Since it was not possible to functionally characterize all alterations found, an *in silico* classification was developed based on 6 bioinformatic programs for missense alterations and 3 programs for splicing alterations. Based on these classification, 435 index cases (175 CI and 260 relatives) had a definite diagnosis of FH, 196 individuals (116 CI and 80 relatives) had a probable diagnosis, 79 a possible diagnosis and 46 were neutral. It was found that 594 subjects had an alteration in the *LDLR* gene, 30 individuals had an alteration in the *APOB* gene and 7 in the *PCSK9* gene.

Another aim of this study was to correlate the genotype with phenotype for all patients referenced to PFHS based on the different molecular diagnostics that were obtained through the classification of the alteration with and without functional studies. Comparison of the phenotypes displayed by children and adults with different types of mutation revealed that children bearing null allele mutations did not present a different phenotype from other mutation type carriers, but the same was not observed for adult patients. This emphasizes that early diagnosis is important for cardiovascular risk stratification of young patients.

The last part of this study focused on exome sequencing of 5 patients with clinical diagnosis of FH but without a mutation identified in genes associated with FH. As exome sequencing produces a large amount of data per sample, a selection of 49 genes associated with lipid metabolism was implemented to restrict the analysis, and 20 alterations distributed over 17 genes were identified. Through this study two alterations in *FLT1* gene in 2 unrelated patients and one alteration in the *SORT1* gene in one patient were identified and found to co-segregate with hypercholesterolemia in the family.

Functional studies of these alterations are required to prove pathogenicity. If any of these alterations is shown to be disease-causing, this will be the first evidence for novel genes causing FH.

Information on the molecular basis of the disease allows correct diagnosis and appropriate therapeutic measures. Genetic diagnosis allows also early identification of relatives at risk, reducing their cardiovascular risk thus leading to an improvement in their quality of life and life expectancy.

**Keywords:** Lipid metabolism, Familial Hypercholesterolemia, functional studies, pathogenicity, pyrosequencing, exome sequencing study.



## ABREVIATURAS

<b>ACAT</b>	acilCoA:colesterol aciltransferase ( <i>acyl-CoA:Cholesterol acyltransferase</i> )	<b>DN</b>	Diagnóstico Neutro
<b>ABCG8</b>	<i>ATP-binding cassette transporter G8</i>	<b>EAM</b>	enfarte agudo do miocárdio
<b>ARH</b>	hipercolesterolemia familiar recessiva	<b>EDTA</b>	ácido etilenodiaminotetraacético
<b>Apo</b>	apolipoproteína	<b>EGF</b>	factor de crescimento da epiderme ( <i>Epidermal Growth Factor</i> )
<b>pb</b>	par de base	<b>EPHF</b>	Estudo Português de Hipercolesterolemia Familiar
<b>cDNA</b>	DNA complementar	<b>ESP</b>	<i>Exome Sequencing Project</i>
<b>CHO</b>	<i>Chinese Hamster Ovary</i>	<b>FACS</b>	<i>Fluorescence Activated Cell Sorter</i>
<b>CI</b>	Caso índice	<b>FDB</b>	deficiência familiar em apoB-100 ( <i>Familial Defective Apolipoprotein B-100</i> )
<b>cLDL</b>	colesterol LDL	<b>FH</b>	hipercolesterolemia familiar
<b>cHDL</b>	colesterol HDL	<b>FITC</b>	isotiocianato de fluoresceína ( <i>fluorescein isothiocyanate</i> )
<b>CYP7A1</b>	colesterol 7 alfa-hidroxilase ( <i>cholesterol 7 alpha-hydroxylase</i> )	<b>g</b>	força centrífuga relativa
<b>DD</b>	diagnóstico definitivo	<b>HDL</b>	lipoproteína de alta densidade ( <i>High Density Lipoprotein</i> )
<b>DCV</b>	doença cardiovascular	<b>HMGCoA</b>	3-hidroxi-3-metilglutaril-coenzima A ( <i>3-hydroxy-3-methylglutaryl coenzyme A</i> )
<b>DGGE</b>	eletroforese em gel de gradiente desnaturante	<b>HRM</b>	<i>High Resolution Melt</i>
<b>dHPLC</b>	cromatografia líquida de alta pressão desnaturante ( <i>Denaturing High Performance Liquid Chromatography</i> )	<b>IDL</b>	lipoproteína de densidade intermédia ( <i>Intermediate Density Lipoprotein</i> )
<b>dL</b>	decilitro ( $10^{-3}$ L)	<b>INSA</b>	Instituto Nacional de Saúde Dr. Ricardo Jorge
<b>DMEM</b>	<i>Dulbecco's Modified Eagle Medium</i>	<b>kb</b>	kilobases
<b>DMSO</b>	dimetilsulfóxido	<b>kDa</b>	kiloDaltons
<b>DNA</b>	ácido desoxirribonucleico	<b>LDL</b>	lipoproteína de baixa densidade ( <i>Low Density Lipoprotein</i> )
<b>dNTP</b>	desoxirribonucleótido trifosfatado	<b>LDLR</b>	receptor das lipoproteínas de baixa densidade
<b>DP</b>	Diagnóstico provável		
<b>DPos</b>	Diagnóstico possível		

<b>LPL</b>	lipoproteína lipase ( <i>lipoprotein lipase</i> )	<b>RT-PCR</b>	<i>reverse transcription polymerase chain reaction</i>
<b>Lp(a)</b>	lipoproteína (a)	<b>SAP</b>	<i>Shrimp Alkaline Phosphatase</i>
<b>LPDS</b>	soro deficiente em lipoproteínas ( <i>lipoprotein deficient serum</i> )	<b>sdLDL</b>	<i>small dense low density lipoprotein</i>
<b>M</b>	concentração molar	<b>SER</b>	<i>sterol regulatory element</i>
<b>mg</b>	miligrama ( $10^{-6}$ g)	<b>SNP</b>	<i>single nucleotide polymorphism</i>
<b>miR</b>	microRNAs	<b>SREBP</b>	<i>sterol regulatory element binding protein</i>
<b>mM</b>	milimolar ( $10^{-3}$ M)	<b>SSCP</b>	polimorfismo de conformação de cadeia única
<b>MLPA</b>	<i>multiplex ligation-dependent probe amplification</i>	<b>TBE</b>	ácido tris/borato/EDTA
<b>mRNA</b>	RNA mensageiro	<b>TG</b>	triglicéridos
<b>NCBI</b>	<i>National Center for Biotechnology Information</i>	<b>U</b>	unidade de actividade de enzima
<b>ng</b>	nanograma ( $10^{-9}$ g)	<b>U937</b>	<i>human leukemia monocyte lymphoma cells</i>
<b>NGS</b>	<i>next generation sequencing</i>	<b>VLDL</b>	lipoproteína de muito baixa densidade ( <i>Very Low Density Lipoprotein</i> )
<b>NHLBI</b>	<i>National Heart, Lung, and Blood Institute</i>	<b>WES</b>	<i>whole-exome-sequencing</i>
<b>PCR</b>	reação de polimerase em cadeia ( <i>Polymerase Chain Reaction</i> )	<b>°C</b>	graus Celsius
<b>PCSK9</b>	<i>proprotein convertase subtilisin/kexin type 9</i>	<b>μL</b>	microlitro ( $10^{-6}$ L)
<b>p/v</b>	peso/volume	<b>pmol</b>	picomol ( $10^{-12}$ mol)
<b>RFLP</b>	<i>restriction fragment length polymorphism</i>		
<b>RNA</b>	ácido ribonucleico		





## CAPÍTULO I

---

### INTRODUÇÃO



# I INTRODUÇÃO

## I.1 Doenças genéticas

A maioria das características dos organismos vivos, como a morfologia, fisiologia, desenvolvimento, comportamento e suscetibilidade a doenças são transmitidas de pais para filhos. As doenças genéticas ocorrem devido a alterações do material genético, mutações, que ocorrem em células germinativas e que são transmissíveis à descendência, ou em células somáticas que não vão ser transmitidas à descendência e cujas consequências para o indivíduo dependem da fase de desenvolvimento em que ocorre a mutação (Lodish *et al.* 2000). As doenças hereditárias são geralmente classificadas em dois grupos com base na sua etiologia: doenças monogénicas ou poligénicas (Kingston 1989).

Numa doença monogénica uma alteração num único gene (como mutações pontuais, deleções ou inserções) é o suficiente para que a doença seja expressa. Estas doenças são muitas vezes referidas como doenças mendelianas uma vez que seguem os padrões de transmissão descritos por Gregor Mendel há mais de 110 anos (Kingston 1989). Estes padrões de transmissão podem ser classificados como autossómica dominante, autossómica recessiva e ligada ao sexo. Para as doenças autossómicas dominantes herdar um alelo defeitivo é o suficiente para que a doença se manifeste. Uma doença com hereditariedade autossómica dominante caracteriza-se por estar presente em todas as gerações e pode ser herdado de qualquer um dos progenitores. Para que haja manifestação da doença basta que um dos progenitores transmita o alelo mutado (Kingston 1989). Exemplos de doenças autossómicas dominantes são: doença de Huntington, anemia falciforme, acondroplasia ou a hipercolesterolemia familiar (FH). No caso das doenças recessivas estas só se manifestam se existirem duas cópias do alelo defeitivo, ou seja, se o indivíduo for homozigótico ou heterozigoto composto. Neste tipo de hereditariedade os pais são apenas portadores da doença. A fibrose cística, talassémia, albinismo e fenilcetonúria são alguns exemplos de doenças que apresentam este tipo de hereditariedade (Kingston 1989).

Para a maioria das doenças monogénicas há uma forte relação entre o genótipo e o fenótipo, o que significa que indivíduos com um dado genótipo relevante para uma doença apresentam uma grande probabilidade de manifestarem o fenótipo característico

associado. A esta probabilidade denomina-se de penetrância. Nos seres humanos existem mais de 5000 doenças mendelianas cuja base molecular é já conhecida e encontram-se registadas na base de dados *Online Mendelian Inheritance in Man* (OMIM) (Online Mendelian Inheritance in Man, 2000).

Na hereditariedade ligada ao sexo, ao contrário da autossómica, cada sexo tem probabilidades diferentes de expressar um alelo. O tipo de hereditariedade mais comum é o recessivo ligado ao cromossoma X. As doenças com este tipo de hereditariedade ocorrem mais frequentemente em homens, dado que estes apresentam apenas um cromossoma X, e são transmitidas através das mães portadoras para o filho afetado. O daltonismo e a hemofilia são dois exemplos bem conhecidos de doenças com hereditariedade recessiva ligada ao cromossoma X (Kingston. 1989).

As doenças poligénicas ou multifactorias (também conhecidas como complexas) apresentam um modo de transmissão que não segue as leis de Mendel. Estas doenças estão associadas a alterações em vários genes em combinação com estilos de vida e fatores ambientais originando fenótipo variáveis. A relação entre um gene particular e a doença não é tão óbvia como nas doenças mendelianas (Badano and Katsanis 2002). São exemplos destas doenças a diabetes, a dislipidemia familiar combinada (FCHL) ou a hipertensão arterial.

## **I.2 Hipercolesterolemia Familiar**

A Hipercolesterolemia Familiar (FH) (OMIM 143890) é umas das doenças mais comum do metabolismo lipídico, apresenta uma hereditariedade autossómica dominante e foi a primeira doença genética do metabolismo lipídico a ser caracterizada molecularmente (Goldstein and Brown 1985). Os doentes com FH apresentam concentrações de colesterol LDL na circulação sanguínea acima do percentil 95 para a idade e sexo, originando doença cardiovascular (DCV) prematura, nomeadamente o enfarte agudo do miocárdio (EAM) (Goldstein *et al.* 1995).

Em 1972 Brown e Goldstein iniciaram os seus estudo com FH homozigotos (HoFH) postulando que seria uma doença que apresentava uma forma autossómica dominante e especulando que o defeito estaria numa proteína que regulava o produto final da síntese do colesterol. Foi um grande desafio para ambos provarem esta teoria, uma vez que

nenhum grupo tinha anteriormente descrito um defeito genético no mecanismo de regulação de *feedback* da síntese do colesterol. (Goldstein *et al.* 1985)

O colesterol e os ácidos gordos realizam funções essenciais quer ao nível da fluidez das membranas e armazenamento de energia como atuando como moléculas sinalizadoras e precursores de outras moléculas (Rifai *et al.* 1999). A maior parte do colesterol presente nos tecidos é sintetizado *de novo* e não absorvido a partir da dieta. O colesterol é produzido em vários tecidos, nos humanos o fígado é o principal local de síntese mas o intestino é também um órgão importante no catabolismo do colesterol. O colesterol é um dos principais componentes das membranas celulares e bainhas de mielina, precursor de hormonas esteroides, vitamina D3 e ácidos biliares e é um dos lípidos com maior relevância clínica (Rifai *et al.* 1999). O colesterol é transportado na corrente sanguínea por lipoproteínas. Nos anos 50 e 60 vários grupos de investigadores (GOFMAN *et al.* 1952; GOFMAN *et al.* 1954; Fredrickson *et al.* 1967) classificaram as lipoproteínas em 4 grandes grupos: lipoproteínas de muito baixa densidade (VLDL, *very low density lipoproteins*), lipoproteínas de densidade intermédia (IDL, *intermediate density lipoproteins*), lipoproteínas de baixa densidade (LDL, *low density lipoproteins*) e lipoproteínas de alta densidade (HDL, *high density lipoproteins*). Estudos feitos posteriormente demonstraram que níveis elevados de partículas LDL no sangue estavam associados a DCV prematura enquanto níveis elevados de HDL mostravam ter um efeito protetor (Goldstein *et al.* 1973). Com base nestas observações vários clínicos começaram a observar quais as partículas que estavam mais presentes no colesterol em circulação dos doentes com FH verificando que eram as partículas de LDL (Fredrickson 1974; Brown and Goldstein 1979).

### **I.2.1 Da observação clínica à descoberta do recetor das LDL**

A FH foi descrita pela primeira vez em 1938 por Carl Müller, um clínico norueguês, que a descreveu como um “erro congénito do metabolismo” que originava níveis elevados de colesterol no sangue e enfarte agudo do miocárdio em indivíduos jovens (Muller 1938) apresentando uma transmissão autossómica dominante. Nos anos seguintes vários grupos (Khachadurian 1964; Kwiterovich *et al.* 1974) continuaram a fundamentar a mesma teoria de hereditariedade mas foram os estudos de Khachadurian em famílias libanesas com FH que elucidaram o modo de transmissão e permitiram dividir a FH em duas formas: heterozigótica e homozigótica (Khachadurian 1964).

Para Goldstein e Brown doentes com HoFH apresentavam o melhor modelo para o estudo da doença uma vez que não tinham influência do alelo normal. Os principais mecanismos de síntese de colesterol ocorrem no fígado e intestino, portanto o ideal seria utilizar um destes tecidos para a realização dos estudos funcionais. Estudos feitos nos finais dos anos 60 por Bailey e Rothblat (Rothblat 1969; Bailey 1973) em linhas celulares de linfoblastos e células L de ratinhos demonstraram que o colesterol era sintetizado nestas células e que esta era uma síntese regulada por *feedback*. Com base nestes estudos, Goldstein e Brown decidiram fazer os seus estudos em fibroblastos de doentes com HoFH de modo a verificar como seria a síntese de colesterol nestas células (Goldstein *et al.* 1974). Em 1973 estes investigadores iniciaram os estudos em fibroblastos de doentes com HoFH e verificaram que a síntese de colesterol era resultado da regulação negativa que ocorre quando nas células. Seguidamente estes investigadores compararam a resposta de fibroblastos de doentes HoFH com indivíduos normais quando a ambas as culturas era adicionado soro, com lipoproteínas, e verificaram que nos indivíduos normais existia uma redução do cLDL enquanto em células de doentes HoFH não. Seguidamente analisaram se o tipo de resposta dos fibroblastos era semelhante quando ambas as culturas celulares eram incubadas (doentes HoFH vs normais) apenas com colesterol e verificam que havia uma redução do colesterol mas não quando incubadas com cLDL apenas. Uma vez que o colesterol parecia conseguir transpor as células por difusão a questão centrava-se como remover o cLDL da circulação para a célula, levando-os a pensar que esta função era desempenhada por um recetor que se ligava ao cLDL e o internalizava para dentro da célula. Esta teoria foi verificada em 1974 quando marcaram o cLDL com iodo radioactivo ( $^{125}\text{I}$ ) e verificaram que nos fibroblastos normais o cLDL era internalizado enquanto de doentes com HoFH o cLDL ficava a superfície (Goldstein and Brown 1974). Em 1982 purificaram a proteína do LDLR (Schneider *et al.* 1982), o cDNA do gene recetor das LDL humano foi clonado pela primeira vez em 1984 (Yamamoto *et al.* 1984) e o isolamento e caracterização do gene foi conseguido em 1985 (Südhof *et al.* 1985). A caracterização molecular da primeira mutação no gene recetor das LDL foi publicada logo de seguida (Goldstein *et al.* 1985). Com esta descoberta estes investigadores foram premiados com o prémio Nobel da medicina em 1985 (Brown and Goldstein 1986).

A forma mais comum da doença é a heterozigótica apresentando uma frequência de 1 em 500 na maioria das populações (Brown and Goldstein 1986). Doentes com FH heterozigótica (HtFH) têm cerca de 50% de recetores de LDL (LDLR) não funcionais originando uma diminuição na remoção do colesterol LDL (cLDL) que desta forma se acumula na corrente sanguínea desde a nascença, originando DCV prematura. A forma homozigótica é rara e severa apresentando uma frequência de 1 em 1 milhão e resulta da herança de 2 alelos mutados, um de cada um dos progenitores. Doentes com FH homozigótica (HoFH) praticamente não produzem recetores funcionais originando valores de cLDL 3 - 5 vezes superiores ao normal e consequentemente desenvolvem DCV na infância (Brown and Goldstein 1986).

### I.2.2 Via do Recetor das LDL

Nos dias de hoje é sabido que as partículas de colesterol das LDL são removidas do plasma por intermédio de uma glicoproteína de superfície específica, o recetor das LDL, que está presente nas membranas celulares e se liga à apolipoproteína B-100 (apoB100) presente na superfície das LDL.

Na membrana celular, os recetores das LDL localizam-se preferencialmente em *coated pits*, regiões especializadas da membrana cobertas por claterina na superfície interior. Quando uma partícula de LDL se liga ao recetor, por reconhecimento da apoB-100 na sua superfície (Figura I:1 (A)), a claterina polimeriza e forma invaginações na membrana que posteriormente se libertam e originam vesículas revestidas de claterina com os complexos gene recetor LDL:LDL no seu interior (Goldstein *et al.* 1995). Outras vezes estes complexos são internalizados com o PCSK9 que se liga a região extracelular do LDLR (Figura I:1 (B)). Para que ocorra a internalização do *LDLR* é necessário, na maioria dos tecidos, uma proteína adaptadora que se liga à cauda citoplasmática do recetor, a proteína adaptadora do recetor das LDL 1 (*LDLRAP1*, *LDL receptor adaptor protein 1*) (Norman *et al.* 1999). As vesículas de claterina que transportam os complexos LDLR:ligando vão fundir-se no citoplasma e formar endossomas onde (Figura I:1 (C)), por diminuição do pH para 6,5 pela ação das bombas de protões ( $H^+$ -ATPases), o recetor vai dissociar-se das partículas de LDL (Davis *et al.* 1986) (Figura I:1 (D)). O recetor é direccionado para vesículas de reciclagem (Figura I:1 (E)), que fazem o seu retorno à membrana celular para realizar novo ciclo e as partículas de LDL vão ser degradadas a nível dos lisossomas, onde a apoB100 é degradada em

péptidos pequenos e aminoácidos. Os ésteres de colesterol existentes no núcleo das partículas das LDL são também hidrolisados, libertando colesterol livre, que é utilizado na síntese de membranas celulares, na produção de hormonas esteróides em tecidos específicos e ácidos biliares (Gillman-Daniel *et al.* 2002) (Figura I.1 (F)). No caso dos recetores que tem o PCSK9 (*proprotein convertase subtilisin/kexin type 9*) na sua superfície a redução do pH aumenta a ligação impossibilitando a dissociação do LDLR e PCSK9, impossibilitando a reciclagem do LDLR (Figura I:1 (D)).

Cada ciclo de atividade do *LDLR* tem a duração aproximada de 10 minutos, ocorrendo internalização e reciclagem mesmo na ausência de ligação das partículas de LDL (Goldstein *et al.* 1995). Possivelmente diferentes proteínas devem interagir com o recetor de modo a direccionar a sua entrada na célula, e embora a maioria destas proteínas não seja ainda conhecida, os seus genes devem ser considerados como possíveis candidatos para a causa hereditária de hipercolesterolemia (Soutar and Naoumova 2007).

A maioria das células têm a capacidade de regular o seu conteúdo em colesterol pois o número de recetores das LDL é regulado pelo seu colesterol livre. A libertação de colesterol livre nos lisossomas leva a: 1) diminuição da síntese endógena de colesterol por inibição da HMG CoA, enzima limitante na biossíntese de colesterol; 2) aumento da taxa de esterificação de colesterol a nível intracelular, reação catalisada pela enzima acil:colesterol acil-transferase (ACAT); e 3) inibição da expressão do gene que codifica o *LDLR* através de elementos de resposta a esteróis (SRE, *Sterol Regulatory Element*) localizados a montante do gene e que são especificamente reconhecidos pelas SREBP (*Sterol Regulatory Element Binding Proteins*) (Goldstein *et al.* 1995; Rifai *et al.* 1999) um fator de transcrição que conduz a expressão dos genes das enzimas envolvidas na síntese do colesterol e do *LDLR*. Desta forma a via do recetor das LDL contribui para manter a homeostase intracelular de colesterol.

#### II.2.2.1 Recetor das LDL: gene e proteína

O gene do recetor das LDL está localizado no braço curto do cromossoma 19, tem aproximadamente 45 kilobases (Kb) e encontra-se dividido em 18 exões e 17 intrões (Figura I:2). O RNA mensageiro (mRNA) transcrito a partir deste gene tem 5.3 Kb, dos quais aproximadamente metade correspondem a uma extensa região 3' não traduzida (UTR) do exão 18 (Goldstein *et al.* 1995). Da tradução do mRNA do LDLR resulta um



polipéptido com 860 aminoácidos. A região a montante do gene *LDLR* contém as sequências reguladoras em *cis* que são responsáveis pela expressão do gene nas células animais. Na região que contém os 200 pares de bases (pb) que antecedem o codão de iniciação da tradução encontram-se três repetições diretas imperfeitas, em que duas das repetições interatuam com o fator de transcrição Sp1 para promover a transcrição e outra que contém um local de regulação condicional positiva por esteróis (SER, *sterol regulatory element*), as TATA boxes e um cluster de locais de iniciação do mRNA (Goldstein *et al.* 1995; Sato and Takano 1995). O *LDLR* é expresso maioritariamente ao nível do fígado, embora seja também expresso em células dos tecidos periféricos. Esta proteína é inicialmente sintetizada sob a forma de um precursor com um peso molecular de 120 kilo Dalton (kDa), que posteriormente migra para o aparelho de Golgi onde sofre glicosilação extensa, atingindo a forma madura com 160 (kDa) (Hobbs *et al.* 1992) (Figura I:1).

Os 18 exões que constituem o gene do recetor das LDL codificam para cinco domínios estruturais da proteína: três extracelulares e dois domínios intracelulares (Figura I:2).

O exão 1 codifica uma região 5' UTR, o péptido de sinal, de 21 aminoácidos hidrofóbicos. Esta região funciona como um sinal para conduzir o recetor dos ribossomas até a membrana do RE. Os exões 2 a 6 codificam 292 aminoácidos do domínio de ligação ao ligando, que possui 7 repetições imperfeitas com aproximadamente 40 aminoácidos cada (repetições I a VII). Várias análises mutacionais verificaram que nem todas as repetições são essenciais para a correta ligação do *LDLR* existindo uma contribuição individual de cada repetição para a capacidade de ligação (Esser and Russell 1988; Goldstein *et al.* 1995). Os exões 7 a 14 codificam para uma sequência de 400 aminoácidos que tem homologia com o precursor do fator de crescimento da epiderme (EGF, *epidermal growth factor*) humano e é constituído por três repetições, A, B (codificam para os exões 7 e 8) e C (codifica o exão 14) ricas em cisteína, intercaladas por uma sequência que contém um motivo conservado Tyr-Trp-Thr-Asp (YWTD) (exões 9-13) e que correspondem ao segundo domínio da proteína, domínio precursor do EGF (Figura I:2). Este domínio é essencial para posicionar o *LDLR* na superfície de modo a ligar-se ao ligando e para a dissociação das lipoproteínas do recetor ao nível dos endossomas quando há um abaixamento do pH (Davis *et al.* 1987). O exão 15 codifica para um terceiro domínio que apresenta 58 aminoácidos ricos em serinas e treoninas que sofrem glicosilação (Figura I:2).

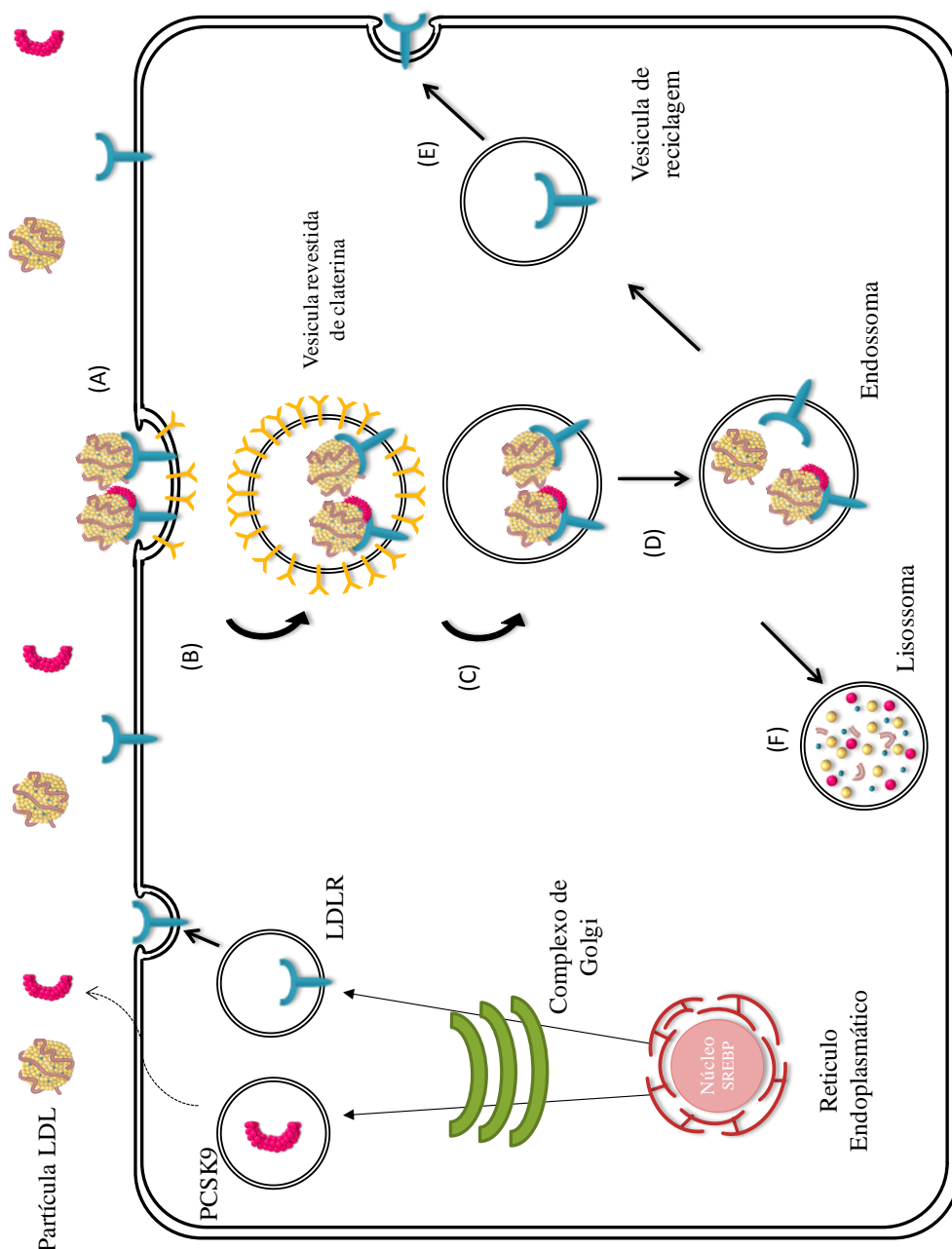


Figura I.1 – Ciclo celular do receptor das LDL. O receptor reconhece as partículas de LDL devido à presença de apolipoproteína B-100 na sua superfície (A) e internaliza na célula numa vesícula de clatrina (B). As vesículas de clatrina fundem-se de forma a formarem endossomas (C). Por diminuição do pH dentro do endossoma ocorre a dissociação dos receptores (D) que são reciclados para a superfície celular (E). As partículas de LDL migram para um lisossoma (F), onde a apolipoproteína B100 é degradada em aminoácidos e o éster de colesterol é convertido em colesterol livre usado pela própria célula no seu metabolismo. Quando o PCSK9 se liga ao LDLR na sua superfície celular os complexos LDLR:LDL:PCSK9 são internalizados em vesículas de clatrina mas já no endossoma e mesmo com a diminuição do pH a ligação LDLR:LDL:PCSK9 aumenta impedindo a dissociação do complexo e reciclagem dos LDLR.

A função do domínio de cadeias de glicolisação tipo O, é ainda desconhecida pensando que possa ser apenas espaçadora (Davis *et al.* 1987). Estudos revelaram que quando este domínio é deletado da proteína não tem qualquer influência na expressão do *LDLR* em fibroblasto (Goldstein *et al.* 1995), contudo um estudo em duas famílias com uma alteração que originada a perda total do exão 15 verificou que esta co-segregava na família com o fenótipo de hipercolesterolemia (Goldstein *et al.* 1995). O exão 16 e a extremidade 5' do exão 17 codificam 22 aminoácidos do domínio transmembranar do *LDLR*. A restante sequência do exão 17 e a região 5' do exão 18 codificam o domínio citoplasmático de 50 aminoácidos que são necessários para a localização do *LDLR* nas vesículas revestidas de claterina e para a endocitose do recetor através do motivo conservado Asp-Pro-Val-Tyr (NPVY) contido nesta cauda citoplasmática (Davis *et al.* 1986; Goldstein *et al.* 1995) (Figura I.2). A observação do domínio extracelular da proteína através da cristalografia proporcionou um modelo interessante de como o recetor das LDL se liga com grande afinidade a membrana da célula libertando o LDLc de seguida no compartimento intracelular apropriado (Rudenko *et al.* 2002).

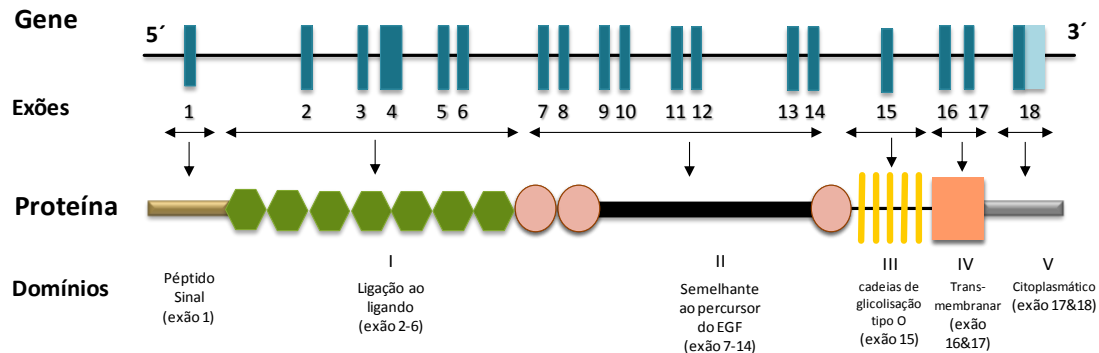


Figura I.2 – Representação esquemática do gene e proteína do recetor das LDL.

#### II.2.2.2 Apolipoproteína B: gene e proteína

Em meados dos anos 80 Grundy e colegas verificaram que a remoção das LDL da circulação era reduzida em alguns doentes com hipercolesterolemia, mas que não tinham um fenótipo claro de FH (Vega and Grundy 1986). Innerarity e colegas verificaram que, em alguns casos, esta redução era causada por uma defeito que diminuía a capacidade das partículas de LDL do doente de se ligarem ao recetor (Innerarity *et al.* 1987). As partículas de LDL ligam-se ao recetor através da apoB (o

seu ligando), levantando a questão de onde estaria o defeito, se na apoB ou nas partículas de LDL. Por ser umas das maiores proteínas nos humanos, polimórfica e com propensão para formar agregados insolúveis em soluções aquosas, assim como sensibilidade à degradação proteolítica, a estrutura primária da apoB apenas foi descrita em 1986 (Chen *et al.* 1986). Após a caracterização do gene da *APOB* foi encontrada uma alteração de um aminoácido no codão 3527 que originava a substituição de uma arginina por uma glutamina e que impossibilitava a ligação das partículas de LDL ao LDLR por parte do ligando das LDL, a apoB (Innerarity *et al.* 1987; Myant 1993). A patogenicidade desta alteração foi efetuada através de estudos em fibroblastos com as partículas de LDL marcadas radioativamente (Innerarity *et al.* 1987) e em células U937 (*human leukemia monocyte lymphoma*) desenvolvidos por Frostegård (Frostegård *et al.* 1990). As células U937 não têm síntese própria de colesterol e dependem da entrega do cLDL exógeno através do recetor LDL. Quando as células são incubadas com soro de um doente com uma mutação no gene *APOB* a sua proliferação é menor quando comparada com a proliferação das células incubadas com soro de um indivíduo normal (Frostegård *et al.* 1990). A doença causada por mutações no gene *APOB* originando um fenótipo de hipercolesterolemia foram denominadas de deficiência familiar em apoB (FDB) (Innerarity *et al.* 1987; Soria *et al.* 1989) no entanto vários autores incluem mutações no gene *APOB* na FH (Goldstein *et al.* 1995; Soutar and Naoumova 2007; Usifo *et al.* 2012).

O gene *APOB* situa-se no braço curto do cromossoma 2 e o mesmo gene produz duas formas de apoB circulante, apoB48 e apoB100. A denominação das proteínas resulta de uma escala em percentil, onde a apoB48 significa que é idêntica em 48% a apoB100 que corresponde ao tamanho total da proteína (Whitfield *et al.* 2004). Nos humanos a apoB48 é produzida no intestino enquanto a apoB100 é produzida no fígado (Powell *et al.* 1987). O gene que codifica para a apoB100 tem 43 kilobases e 29 exões, sendo o exão 26 o maior com 7572 pb, codificando mais de metade da proteína total que apresenta 4563 aminoácidos (Innerarity *et al.* 1987; Myant 1993).

A função da APOB está relacionada com a sua estrutura, ela serve para manter a integridade das partículas de LDL assim como para controlar os níveis plasmáticos de colesterol LDL na circulação através da sua ligação ao recetor das LDL (Brown and Goldstein 1986). As partículas de LDL apresentam à superfície apoB100 que rodeia toda a partícula numa forma semelhante a uma corda dando forma e tamanho às

partículas assim como afinidade para a ligação das partículas ao recetor das LDL (Miserez and Keller 1995). A região composta pelos aminoácidos 3386 ao 3396 é descrita como sendo a zona de ligação ao LDLR (März *et al.* 1993), no entanto a mutação mais comum no gene *APOB* encontra-se descrita fora desta região, p.Arg3527 (Innerarity *et al.* 1987). Quando descreveram o mecanismo molecular da FDB Boren e colegas (Boren *et al.* 2001) verificaram que alterações nos resíduos (3174 a 3184 e 4181 a 4540) destabilizam *clusters* importantes na conformação da apoB levando a que o LDLR não reconheça as partículas de LDL (Chatterton *et al.* 1995; Boren *et al.* 2001). Boren e seus colegas demonstraram que as alterações p.Arg3507Trf, p.Arg3527Gln, p.Arg3558Cys, p.Trf4396Tyr são importantes para a conformação da proteína apoB100 e que uma alteração num destes aminoácidos altera a sua conformação impedindo a ligação entre a partícula de LDL e o LDLR. Em 2007, Krisko e Etchebest propuseram um modelo para a estrutura terciária da apoB definindo 8 domínios estruturais, em contraste com Segrest que a caracterizou em 5 domínios estruturais (Segrest *et al.* 2001), e estabelecendo uma nova região de ligação ao recetor situada entre os aminoácidos 2820-3202 e 3243-3498. No entanto este modelo teórico para a estrutura terciária da proteína apoB não foi comprovado experimentalmente uma vez que é difícil obter-se a estrutura cristalizada da APOB por esta ser parte integrante da partícula LDL e nunca se encontrar no meio dissociada destas partículas (Prassl and Laggner 2009).

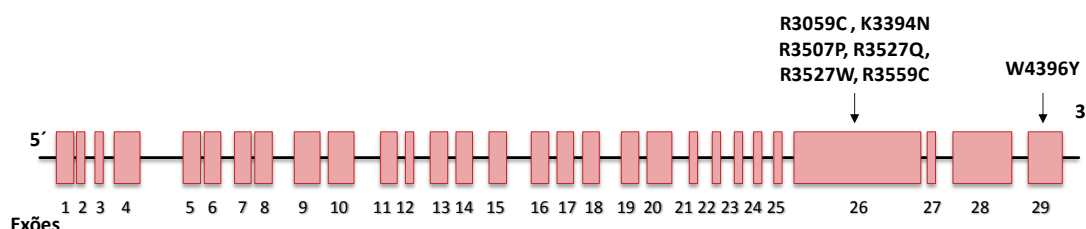
Foram descritas outras alterações no gene *APOB* associadas a hipercolesterolemia, p.Asn1914Ser, p.His1923Arg, e p.Ala4481Thr (Meng *et al.* 1996). Estudos funcionais a estas alterações demonstraram estas são polimorfismos que não afetam a função normal da proteína (Meng *et al.* 1996).

Ao contrário do que acontece no *LDLR* onde há uma grande heterogeneidade de mutações, na apoB100 estão apenas descritas algumas variantes, sendo a alteração p.Arg3527Gln a mais comum nos indivíduos com descendência europeia (Innerarity *et al.* 1987; Soria *et al.* 1989). A frequência desta mutação na Europa é 1/1000 (Myant 1993), mas em alguns países esta frequência é mais elevada, como é o caso da Polónia que apresenta uma frequência de 1/250 (Bednarska-Makaruk *et al.* 2001). Nos europeus pensa-se que esta mutação tem origem num haplótipo raro de um ancestral comum que terá vivido há cerca de 6750 anos. A penetrância desta mutação não é de 100% justificando o facto de doentes com mutações na *APOB* apresentarem um fenótipo menos severo que os doentes HtFH (Myant 1993; Vrablík *et al.* 2001). Outra mutação

no mesmo codão mas que resulta da substituição da arginina pelo triptofano (p.Arg3527Trp) é comum em doentes hipercolesterolémicos na China e Malária mas rara na Europeia

Em 2012 foram descritas duas novas alterações na apoB, (Motazacker *et al.* 2012) como causa genética de FDB em duas famílias estudadas por *exome sequencing*. Estudos funcionais em linhas celulares de HepG2 demonstraram que estas alterações originam uma redução da ligação da partícula de LDL ao seu recetor (Motazacker *et al.* 2012). Também em 2012 um estudo por *exome sequencing* efetuado a 48 doentes ingleses com FH detetou uma alteração na *APOB*, uma substituição de uma arginina por uma valina na posição 3426, no entanto neste caso não foram feitos estudos funcionais a demonstrar a patogenicidade da alteração (Futema *et al.* 2012) (Figura I.3).

A



B

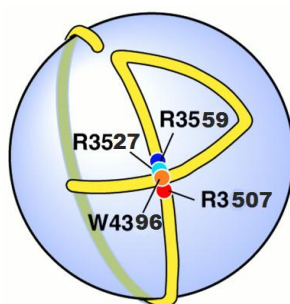


Figura I.3 – A) representação do gene da APOB. B) Modelo da ligação da apoB ao recetor das LDL. A ligação do recetor à apoB depende da interação entre os resíduos de arginina 3507, arginina 3527, arginina 3558 e o resíduo de triptofano 4396. Esta interação é essencial para o correcto enrolamento da extremidade carboxi-terminal, permitindo a ligação entre a partícula de LDL e o seu recetor. (Adaptado de Boren *et al.* 2001)

### II.2.2.3 PCSK9: gene à proteína

Em 2003, após estudos de *linkage* em duas famílias francesas, um novo gene foi associado à hipercolesterolemia familiar, o gene *PCSK9* que codifica para a proteína pró-proteína convertase subtilisina/quexina tipo 9 (PCSK9, *proprotein convertase*

*subtilisin/kexin type 9*) (Abifadel *et al.* 2003). Mutações neste gene co-segregam com uma hipercolesterolemia severa na maioria dos doentes identificados em todos os países, apresentando uma transmissão autossómica dominante. O facto de ser um gene de expressão regulada por esteróides apoia o seu envolvimento no metabolismo do colesterol (Maxwell *et al.* 2003; Horton *et al.* 2007).

O PCSK9 situa-se no braço curto do cromossoma 1 (1p34.1-p32), abrange 3617 pb ao longo de 12 exões e codifica para uma glicoproteína de 692 aminoácido. Anteriormente identificada como neural apoptose convertase regulada (*neural apoptosis regulated convertase 1*, NARC-1) esta proteína foi caracterizada por Seidah e colegas como membro da subfamília das pró-proteína convertases (Seidah *et al.* 2003). A proteína é composta por um péptido de sinal, um pro-domínio, um domínio catalítico e um domínio C-terminal (Figura I:5) sendo produzida essencialmente no fígado e intestino (Seidah *et al.* 2003). A proteína é sintetizada como uma pró-enzima inativa que contém três resíduos (p.Asp186, p.His226 e p.Ser386) necessários para atividade catalítica (Seidah *et al.* 2003). A forma precursora do PCSK9 apresenta um tamanho de 74 kDa e sofre uma clivagem autocatalítica no retículo endoplasmático que produz um fragmento catalítico de 60 kDa. No RE sofre uma clivagem autocatalítica do zimogénio, essencial para o transporte deste compartimento e para a secreção (Benjannet *et al.* 2004; McNutt *et al.* 2007). O prodomínio clivado de 14 kDa permanece associado com o domínio catalítico permitindo que a proteína madura se mova do ER até a via secretora onde ocorrerá o enrolamento correto da proteína, bem como, a regulação da atividade catalítica da enzima interagindo com o LDLR (Benjannet *et al.* 2004; McNutt *et al.* 2007).

O PCSK9 revelou ser um gene bastante polimórfico influenciando os níveis de colesterol plasmático quer em indivíduos normolipidémicos bem como em doentes com FH. Alterações neste gene originam fenótipos antagónicos, uma vez que foram descritas alterações como responsáveis por níveis reduzidos de colesterol (em normolipidémicos) assim como níveis elevados de cLDL em doentes com FH (Abifadel *et al.* 2009), (Figura I:5). Estudos efetuados por Maxwell em fígados de ratinhos *wild type* demonstraram que a sobre expressão do gene PCSK9 leva a um aumento de colesterol para o dobro, pela diminuição da quantidade e funcionalidade da proteína LDLR, mas sem alteração dos níveis de mRNA do LDLR (Maxwell and Breslow 2004). Por outro lado estudos em fígados de ratos transgénicos, sem o gene PCSK9, revelaram existir um

aumento nos níveis de proteína LDLR e consequentemente uma diminuição nos níveis de colesterol LDL (Rashid *et al.* 2005). Assim pode-se considerar que alterações com ganho de função originam hipercolesterolemia autossômica dominante (Abifadel *et al.* 2003; Naoumova *et al.* 2005; Horton *et al.* 2007) enquanto mutações com perda de função do *PCSK9* originam um aumento do número de *LDLR* no hepatócito e consequentemente no plasma, levando a uma redução do colesterol LDL (Cohen *et al.* 2005; Berge *et al.* 2006).

A função da proteína *PCSK9* no ciclo do recetor das LDL tem sido amplamente investigada tendo sido proposto um mecanismo de interação entre as proteínas (Horton 2009). O prodomínio clivado permanece associado com o domínio catalítico permitindo que a proteína madura seja segregada pela célula. Uma vez na circulação o *PCSK9* apresenta uma grande afinidade para se ligar à região extracelular do *LDLR* sendo internalizado com o recetor, na maioria das vezes através das vesículas de claterina (Figura I:1 (B)). Aparentemente o *LDLR* parece não competir com o *PCSK9* para que ocorra a formação dos complexos *LDLR:LDL* (Soutar 2010). Já no endossoma, e mesmo ocorrendo a redução do pH, a ligação *PCSK9:LDLR* aumenta não havendo dissociação do complexo (Figura I:1 (D)) e impossibilitando desta forma que o recetor seja reciclado em vesículas até a superfície da célula para que faça a sua função de recolha do cLDL (Cunningham *et al.* 2007), por outro lado as partículas de LDL não vão ser degradadas ao nível dos lisossomas (Soutar 2010) (Figura I:1 (F)). Um dos aspetos mais interessantes do *PCSK9* é o facto da atividade catalítica da convertase parecer estar restringida a sua autoclivagem, ao passo que a maioria das pró-proteínas convertases são ativadas como proteases por outros substratos de forma a ocorrer a autoclivagem. No caso do *PCSK9* estudos estruturais revelaram que os blocos do prodomínio ligados e clivados interagem com o centro ativo (Cunningham *et al.* 2007). A evidência científica para este facto é a substituição de um resíduo de serina, no centro ativo da proteína *PCSK9*, por um resíduo de alanina que impede qualquer autoclivagem, mas quando estes clones de *PCSK9* foram co-expressos em cultura celular, esta substituição não pareceu afetar a capacidade do *PCSK9* para reduzir os níveis de recetor das LDL (McNutt *et al.* 2007).



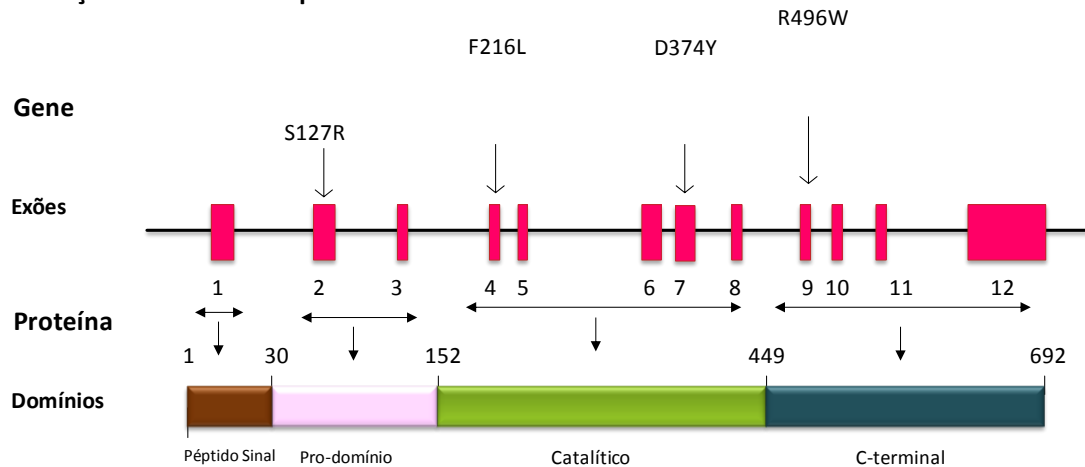
**Mutações associadas a hipercolesteromia familiar**

Figura I.4 – Representação esquemática do gene e proteína do PCSK9.

As mutações associadas a ganho-de-função do PCSK9 são raras, mas estão na sua maioria caracterizadas funcionalmente. Em Portugal, até ao momento, apenas foi encontrada uma mutação em dois indivíduos não relacionados (Bourbon *et al.* 2008). As alterações que originam perda-de-função do PCSK9 apresentam uma maior frequência em todas as populações, no entanto as mutações que levam a produção de um alelo nulo estão associadas a penas ao fenótipo de hipocolesterolemia maioritariamente em indivíduos afro-americanos (Cohen *et al.* 2005).

#### II.2.2.4 Proteína adaptadora do recetor das LDL 1 (LDLRAP1) -Hipercolesterolemia autossómica recessiva

Em 1964 quando Khachadurian estudava famílias libanesas com FH constatou que um número reduzido de famílias, com um fenótipo severo de FH, apresentava um padrão de transmissão recessiva ao contrário da maioria das famílias onde esta era dominante (Khachadurian 1964). A mesma observação foi reportada em famílias da Sardenha anos mais tarde (Zuliani *et al.* 1999). Contudo estudos funcionais em fibroblastos não detetaram nenhuma alteração genética em famílias com uma transmissão recessiva (Soutar 2010). Estudos em linfoblastos de doentes com fenótipo de FH recessiva foram desenvolvidos revelando que a expressão do *LDLR* à superfície era normal mas que este não era internalizado (Norman *et al.* 1999). Em 2001 Garcia e colegas descreveram pela primeira vez a forma recessiva da hipercolesterolemia familiar (ARH) como sendo causada por mutações na proteína adaptadora do recetor das LDL 1 (LDLRAP1) (Garcia *et al.* 2001) após estudos de *linkage* em famílias libanesas e da sardenha.

Estudos em linfoblastos de doentes com alterações neste gene revelaram afetar a internalização do complexo LDLR:LDL. A proteína LDLRAP1 contém um domínio de ligação fosfotirosina-conservado com 130 aminoácidos que está presente noutras proteínas adaptadoras que se ligam a uma sequência motivo NPXY na cauda citoplasmática de vários recetores da superfície celular, incluindo o recetor das LDL. A porção C terminal da LDLRAP1 contém uma região de consenso estruturada de blocos de claterina que se liga a uma cadeia conservada de claterina com 27 aminoácidos unindo-se à subunidade beta-2-adaptina do adaptador da proteína (AP)-2, um componente estrutural das vesículas de claterina (He *et al.* 2002). Estes dados sugerem que a função da LDLRAP1 é interagir com o *LDLR*, através do domínio citoplasmático, permitindo que toda a maquinaria de vesículas revestidas de claterina se forme e ocorra a endocitose (Soutar and Naoumova 2007). No entanto o mecanismo de endocitose destas vesículas de claterina não é totalmente conhecido sabendo-se apenas que a proteína LDLRAP1 apresenta mecanismos estruturais compensatórios para promover a sua ligação ao *LDLR* (Dvir *et al.* 2012) (Figura I.4). Assim a LDLRAP1 parece ser a proteína responsável nos hepatócitos pela endocitose do complexo LDLR:LDL (Mishra *et al.* 2002; Mettlen *et al.* 2010), no entanto estudo em fibroblasto de doentes com as mesmas mutações revelaram não afetar o processo de endocitose (Garcia *et al.* 2001; Arca *et al.* 2002). É sabido que em certos tipos de células outras proteínas adaptadoras podem substituir a LDLRAP1, uma das proteínas candidatas é a *Disable-2* (Dab2) que apresenta uma organização estrutural e de domínios muito semelhante (Eden *et al.* 2007) a LDLRAP1. No entanto não foi ainda demonstrado que ocorra uma substituição da proteína Dab2 pela LDLRAP1 nos fibroblastos de doentes com ARH (Mettlen *et al.* 2010; Soutar 2010).

O gene *LDLRAP1* está localizado no braço curto do cromossoma 1 (1p35) e é composto por 9 exões, o cDNA tem 2940 pb e codifica para uma proteína com 308 aminoácidos (Garcia *et al.* 2001). Até a data encontram-se descritas poucas alterações no gene *LDLRPAI*, todas as mutações originam um alelo nulo. Na Sardenha a ARH é comum, presumivelmente devido a um efeito fundador naquela região (Arca *et al.* 2002). Habitualmente os pais dos indivíduos afetados são heterozigóticos e têm valores normais de colesterol, mas são necessários mais estudos de investigação para determinar o risco DCV a longo prazo desta população (Soutar *et al.* 2003).

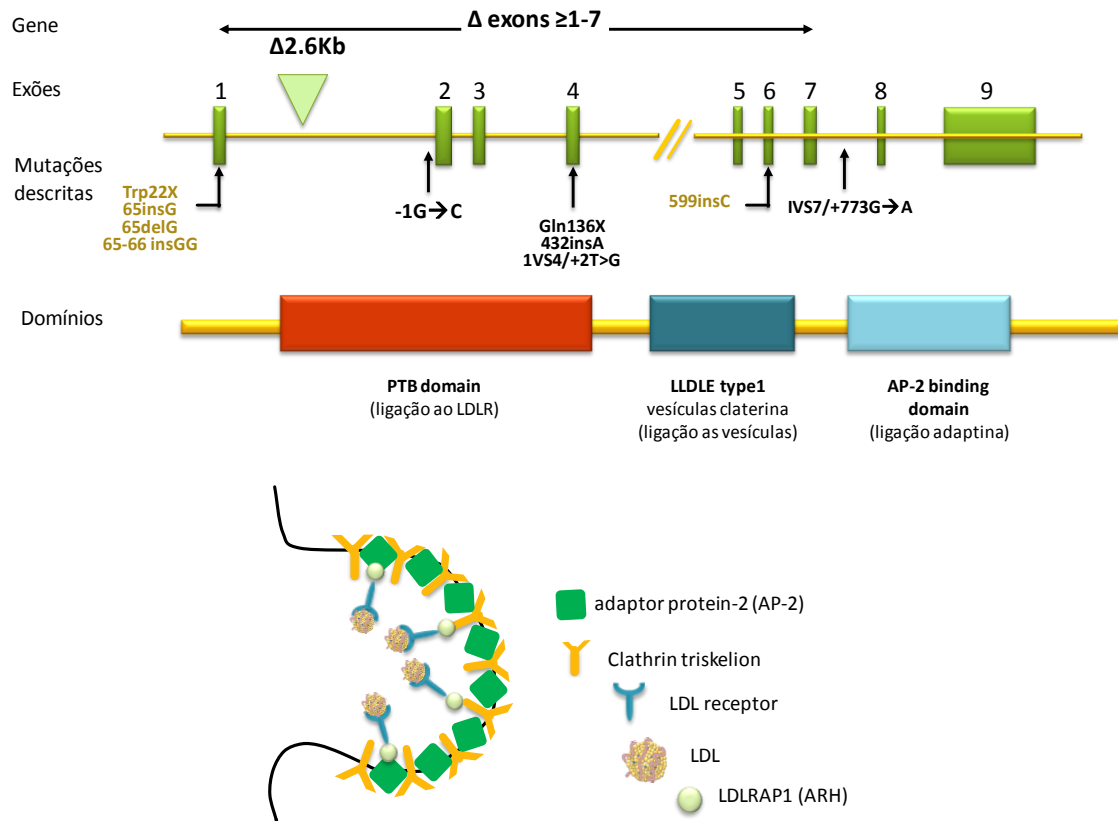


Figura I.5 – Representação gráfica do gene e proteína do LDLRAP1.

(Adaptado de Soutar and Naoumova 2007)

O fenótipo da ARH é muito semelhante ao dos doentes com HoFH, no entanto apresentam valores médios de colesterol total e colesterol LDL mais baixos e respondem melhor ao tratamento por fármacos e LDL-aferese (Soutar *et al.* 2003). Como referido anteriormente a ARH é a forma recessiva da doença, no entanto não deve ser descartada a hipótese de algumas mutações resultarem num fenótipo dominante. As interações gene-gene e gene-ambiente também podem influenciar a função do recetor das LDL e, consequentemente, os níveis plasmáticos em alguns doentes (Soutar *et al.* 2003). Por estas razões, a herança de um só alelo *LDLRAP1* mutado também deve ser considerada, quando se estudam doentes com hipercolesterolemia não justificada por outras formas.

### I.2.3 Outras causas de hipercolesterolemia

Apesar de todos os avanços tecnológicos nos métodos de pesquisa de alterações, existe um número de indivíduos com fenótipo clínico de FH e sem mutações num dos genes

associados a FH. Nestes indivíduos a causa genética pode estar noutro gene, mas estes casos são raros.

Até ao momento foram descritas alterações em 2 genes envolvidos no metabolismo do lipídico, que originam um fenótipo de hipercolesterolemia familiar: colesterol 7 alfa-hidroxilase (CYP7A1, cholesterol 7 alpha-hydroxylase) e proteínas de ligação ao elemento de resposta a esteróis 2 (SREBP-2, sterol regulatory element binding protein 2) (Soutar and Naoumova 2007).

O gene CYP7A1 é expresso unicamente no fígado e a enzima CYP7A1 catalisa o primeiro passo do catabolismo hepático do colesterol e a síntese de ácidos biliares. O sistema hepato-biliar é a primeira via pela qual o colesterol é excretado. A síntese dos ácidos biliares tem um papel importante na manutenção da homeostase do colesterol nos mamíferos, uma alteração na região promotora do CYP7A1 foi descrita anteriormente como originando um fenótipo de sensibilidade de um indivíduo à composição da sua dieta. Pullinger e colaboradores estudaram três indivíduos com um fenótipo recessivo de hipercolesterolemia e identificaram uma alteração que origina um codão *stop* prematuro e por sua vez uma proteína truncada originando uma enzima não funcional (Pullinger *et al.* 2002). Os níveis de colesterol LDL destes 3 doentes eram semelhantes aos apresentados por doentes com HoFH e apresentavam litíase biliar prematura. O estudo familiar revelou que os indivíduos heterozigotos para esta alteração apresentavam hipercolesterolemia indicando que esta alteração tinha uma transmissão autossómica recessiva (Pullinger *et al.* 2002). Pensa-se que uma deficiência no CYP7A1 possa causar um decréscimo de produção de ácidos biliares e uma acumulação de colesterol no fígado, levando a uma sub-regulação dos recetores das LDL (Pullinger *et al.* 2002; Soutar and Naoumova 2007).

O gene CYP7A1 está localizado no cromossoma 8q11-q12 e é composto por 6 exões. Até hoje foi descrita apenas uma mutação como originando FH (Pullinger *et al.* 2002).

O SREBP-2 tem um papel chave na regulação do metabolismo de ácidos gordos e colesterol. O gene SREBP-2, localiza-se no cromossoma 22, compreende 72 kb e possui 19 exões (Hua *et al.* 1995; Miserez *et al.* 2002). O SREBP é um fator de transcrição da família leucina hélice-loop-hélice (bHLH-zip) que contém sequências acídicas no seu terminal seguidas por uma sequência rica em leucinas e prolinas e por uma extensão de 120 aminoácidos ricas em resíduos de glutamina (Hua *et al.* 1993). O domínio ácido é necessário para a ativação da transcrição. A jusante deste domínio existe o domínio

hélice-loop-hélice (bHLH) que medeia a dimerização e ligação do SREBP ao DNA. Logo após o domínio bHLH-zip surge uma extensão de resíduos hidrofóbicos que se liga a membrana do RE e complexo de Golgi originando uma orientação sob a forma de gancho (Hua *et al.* 1993). O nível de colesterol nas células é controlado pela regulação da quantidade de colesterol sintetizado e/ou removido a partir da corrente sanguínea através do recetor de LDL. Quando as células estão privadas de colesterol, o SREBP sofre uma clivagem proteolítica que origina a libertação de metade do terminal amino da proteína da membrana. Esta forma truncada e madura do SREBP, que contém a ativação da transcrição e o domínio bHLH zip, entra no núcleo ligando-se uma sequência de 10 pb, denominada elemento de resposta a esteróis 1 (SRE1) na região promotora de genes do metabolismo do colesterol, incluindo o recetor de LDL (Horton and Shimomura 1999). Desta forma promove a transcrição do gene e a expressão do recetor de LDL originando a absorção de colesterol LDL do plasma. Um polimorfismo no domínio regulador do *SREBP2* revelou estar envolvido numa redução da transcrição do gene recetor de LDL, o que sugere que alterações neste gene podem causar hipercolesterolemia (Miserez *et al.* 2002). Foram também descritas algumas alterações nestes genes como sendo a causa da hipercolesterolemia em doentes com fenótipo de FH (Muller and Miserez 2002). No entanto a evidência que estas alterações são a justificação para o fenótipo destes doentes não é forte nem conclusiva (Soutar and Naoumova 2007).

#### **I.2.4 Caracterização molecular da FH**

As mutações no LDLR são bastante heterogéneas por todo o mundo, exceto em algumas regiões onde se verifica a presença de um efeito fundador, como é o caso dos judeus *Askenazi* descendentes da Lituânia, franco-canadianos, africanos, hebreus, finlandeses e cristãos libaneses (Hobbs *et al.* 1992). Até ao momento estão descritas e publicadas na base de dados da FH (<http://www.ucl.ac.uk/ldlr/Current/>; <https://grenada.lumc.nl/LOVD2/UCL-Heart/home.php>) mais de 1300 mutações apesar de a grande maioria não ter sido caracterizada funcionalmente. As primeiras mutações descritas na FH foram grandes deleções uma vez que a técnica utilizada para a pesquisa de alterações no *LDLR* era o *Southern blotting*, uma técnica que permite a deteção de moléculas de DNA, após a sua separação electroforética, por hibridação com uma sonda, cuja sequência é complementar à desse DNA. Com o surgimento de novas tecnológicas que permitiram

amplificação do DNA por reação de PCR (polymerase chain reaction) e sequenciação por Sanger, foi possível identificar mutações pontuais e pequenas deleções/inserções aumentando de forma exponencial as mutações deste tipo (Soutar and Naoumova 2007). Desta forma vários tipos de mutações foram descritos em doentes com FH, incluindo mutações que resultam da substituição de um único nucleótido levando à formação de um codão que codifica um aminoácido diferente do original (mutações *missense*); mutações que resultam da substituição de um nucleótido que leva a formação de um codão diferente continuando no entanto a codificar o mesmo aminoácido (mutações silenciosas); mutações que resultam da substituição de um nucleótido que leva à formação de um codão *stop* prematuro e consequentemente a uma proteína truncada (mutações *nonsense*); mutações causadas por pequenas deleções ou inserções que provocam um desfasamento na grelha de leitura originando uma proteína truncada (mutações *frameshift*); mutações na região do promotor e mutações que afetam o *splicing* correto do pré-mRNA (mutações de *splicing*). A análise molecular das mutações no *LDLR* levou a vários resultados importantes que permitiu um melhor conhecimento sobre a função dos diferentes domínios do recetor das LDL (Soutar and Naoumova 2007). Para alguns tipos de mutações, como é o caso dos grandes rearranjos, mutações *nonsense* e mutações *frameshift* a patogenicidade da mutação não é usualmente contestada, assumindo-se que será a causa genética da doença. Em contraste, mutações pontuais e mutações que afetam o *splicing* correcto do RNA são mais difíceis de aferir quanto a sua patogenicidade. A maioria das alterações detetadas no *LDLR* que levam a substituição de um único aminoácido tem um efeito deletério na atividade do recetor. No entanto, depois da realização de estudos funcionais foi provado que algumas destas alterações descritas como causadores de FH que não são patogénicas (Lombardi *et al.* 1997; Naoumova *et al.* 2004; Etxebarria *et al.* 2012; Silva *et al.* 2012). Idealmente as mutações pontuais deveriam ser estudadas em células heterólogas *in vitro* e a sua atividade comparada com a atividade normal do *LDLR*. Para uma mutação poder ser considerada patogénica tem de satisfazer os seguintes critérios: co-segregação da alteração com a hipercolesterolemia na família; ausência da alteração num grande grupo de indivíduos normolipidémicos ( $\geq 100$ ); conservação do aminoácido afetado em várias espécies a alteração da natureza da substituição do aminoácido afetado; a natureza do aminoácido afetado e estudos funcionais (Cotton and Scriver 1998). Para as alterações que afetam o *splicing* existem programas que fazem a previsão *in silico* e permitem prever se a alteração afetará ou não o *splicing* correto, sendo no

entanto falíveis, mas apenas os estudos funcionais por transcriptase reversa do mRNA (RT-PCR) do indivíduos permite confirmar se a alteração é ou não patogénica. Alterações sinónimas (ocorre uma mudança de uma base, mas não há mudança de aminoácido) são usualmente consideradas como não patogénicas, no entanto foram já descritas algumas alterações sinónimas que introduzem uma nova região consenso de *splicing* originando uma mutação patogénica (Bourbon *et al.* 2007).

Ao longo dos últimos 30 anos dois tipos de ensaios têm sido utilizados para comprovar a patogenicidade das alterações no *LDLR* em doentes com FH. O primeiro método, desenvolvido por Brown e Goldstein, baseia-se na determinação dos níveis de ligação, internalização e degradação de LDL marcado radiativamente com  $^{125}\text{I}$  na porção proteica do LDL e eram inicialmente realizados em fibroblastos e linfoblastos de doentes (Goldstein and Brown 1979; Knight *et al.* 1987). Através deste método foi possível estudar o efeito das diferentes mutações detetadas ao longo do gene *LDLR* demonstrando que afetam diferentes fases do ciclo do recetor (Brown and Goldstein 1986). O segundo método foi desenvolvido mais recentemente e utiliza a citometria de fluxo para determinar a taxa de ligação e internalização de LDL marcado com 3,3'-di-octadecylindocarbocyanine (DiI), um fluoróforo lipofílicos, ou com recurso a anticorpos específicos para o *LDLR*, marcados secundariamente com fluorescência, de forma a avaliar os níveis de expressão do recetor à superfície (Jensen *et al.* 1994; Etxebarria *et al.* 2012). Com a utilização deste tipo de marcação é também possível analisar a expressão do recetor por microscopia de fluorescência. Vários estudos já demonstraram que este método é fiável para caracterizar mutações no recetor das LDL (Raungaard *et al.* 1998; Raungaard *et al.* 1999; Etxebarria *et al.* 2012). Os estudos funcionais são importantes para comprovarem se uma alteração é ou não patogénica, no entanto os estudos anteriormente foram realizados utilizando células do doente não podendo garantir que não ocorra interferência de outros fatores genéticos na modulação do efeito das alterações, para além do facto de existir interferência do alelo normal no caso de estudos realizados com células de doentes heterozigóticos. De forma a se ultrapassar este problema, desenvolveram-se sistemas heterólogos de expressão *in vitro* que permitem o estudo do efeito do alelo mutado na função do *LDLR*, sem interferência do alelo normal ou de combinações de fatores genéticos dos doentes (Hobbs *et al.* 1992; Silva *et al.* 2012).

O diagnóstico molecular de FH baseia-se na identificação de mutações em três genes que foram associados ao fenótipo de FH: *LDLR*, *APOB* e *PCSK9*. Mutações no gene *LDLR* são a causa mais comum de FH, cerca de 60-80% dos doentes têm uma mutação neste gene (Tosi *et al.* 2007). Mutações no gene *APOB* têm sido descritas em cerca de 2 a 10%, dependendo das populações (Humphries *et al.* 2008) e mutações no gene *PCSK9* (proprotein convertase subtilisin/kexin type 9) são uma causa rara de FH, só cerca de 1% dos doentes com FH tem mutações neste gene (Soutar and Naoumova 2007). No entanto continua por explicar o fenótipo de cerca de 40-50% dos indivíduos com diagnóstico clínico de FH.

### **I.2.5 Tratamento**

Doentes com FH apresentam um elevado risco para desenvolverem DCV, não só por apresentarem níveis de colesterol plasmático muito elevados, mas também porque estão expostos a níveis elevados desde a infância (Bourbon *et al.* 2006; Perk *et al.* 2012). De forma a reduzir os níveis de colesterol LDL no plasma foram desenvolvidos diferentes tipos de fármacos que atuam nas diferentes vias metabólicas restabelecendo os níveis normais das partículas de LDL. Esta redução pode ser conseguida através: da inibição da produção de colesterol endógeno; aumento da síntese da excreção de ácidos biliares; ou da inibição da absorção intestinal de colesterol. Os inibidores da redutase da 3-hidroxi-3-metilglutaril-coenzima A (estatinas) são o fármaco de primeira linha na hipercolesterolemia familiar (Catapano *et al.* 2011), mas pode haver necessidade de recorrer a outros fármacos, com mecanismos de ação específicos e por vezes complementares. Nas crianças estão recomendadas principalmente as resinas, desde os 2 anos de idade, e as estatinas, a partir dos 8 anos, sempre que o risco global o justifique. Crianças e adultos com FH devem ser seguidos em consulta especializada (Goldberg *et al.* 2011).

O diagnóstico atempado da FH e a introdução de terapêutica adequada precocemente podem reduzir significativamente o número de eventos cardiovasculares e evitar mortes prematuras, o que tem custos familiares e sociais muito importantes.



### **I.2.6 Estudo Português de Hipercolesterolemia Familiar**

O estudo Português de Hipercolesterolemia Familiar (EPHF) teve início em 1999; até então a identificação e caracterização molecular de doentes com FH em Portugal era praticamente inexistente (Bourbon and Rato 2006). O EPHF tem como objetivo principal a realização de um estudo epidemiológico para a determinação da prevalência e distribuição da Hipercolesterolemia Familiar em Portugal, bem como a caracterização clínica e molecular dos doentes. Tem também como objetivo secundário a caracterização da patofisiologia da DCV em doentes com FH (Bourbon and Rato 2006). O diagnóstico molecular da FH permite a correta identificação da patologia e fundamenta a instituição de terapêutica farmacológica mais agressiva e/ou precoce, com a consequente redução do risco cardiovascular nos indivíduos afetados (Bourbon and Rato 2006; Bourbon *et al.* 2008). O estudo molecular identificou até à data 550 indivíduos com FH, que se encontram de momento a receber aconselhamento e tratamento de acordo com a sua condição. No entanto a causa genética em mais de 300 indivíduos portugueses com diagnóstico clínico de FH continua por explicar. Apesar do EPHF existir há mais de 13 anos a FH está sub-diagnosticada no nosso País. No entanto, por falta de identificação clínica destes doentes, estima-se que existam cerca de 20.000 casos de FH (Bourbon and Rato 2006; Bourbon *et al.* 2008).

### **I.3 Mapeamento genético**

Nos últimos 30 anos diversas técnicas de biologia molecular foram desenvolvidas de modo a mapear o genoma humano e dar uma maior informação e conhecimento sobre doenças monogénicas e poligénicas. O projeto de sequenciação do genoma humano foi um ponto de viragem para algumas técnicas de mapeamento genético e assistiu-se a um avanço tecnológico ainda maior após este projeto, bem como um conhecimento mais aprofundado de algumas doenças (Collins *et al.* 2003) (Frazier *et al.* 2003) (Human Genome Sequencing Consortium International 2004).

O método de sequenciação de Sanger foi a metodologia preferencial para sequenciação de ácido desoxirribonucleico (DNA) nos últimos 30 anos. Era um dos métodos preferenciais para a pesquisa de mutações em doenças mendelianas, mas não o único. Na década de 80 surgiram os primeiros estudos de *linkage*. A análise de *linkage* baseia-se na ocorrência de *crossing-over* ou recombinação durante a meiose. A frequência de

recombinação entre dois *loci* depende da distância física e genética entre os dois. Quanto menos afastados estiverem dois *loci* num cromossoma, menor será a probabilidade de serem separados por um *crossing-over*, pelo que alelos num pequeno segmento cromossómico tendem a ser transmitidos em bloco (diz-se que os dois *loci* estão em *linkage*). A fração de recombinação (teta) (proporção de meioses em que ocorre recombinação) é uma indicação da distância genética entre dois *loci* (Burton *et al.* 2005).

Se numa família um marcador segregar consistentemente com a doença, isso implica a existência de um locus de susceptibilidade para a doença na região do marcador (Dawn Teare and Barrett 2005). No entanto esta técnica apresenta algumas limitações, uma delas é ser necessário ter um número significativo de familiares de várias gerações afetados e não afetados para que exista um bom poder estatístico; outra limitação é o facto de os estudos de *linkage* apenas apontarem uma região no genoma candidata para a doença, sendo necessário sequenciar todos os genes nessa região, tornando o trabalho moroso e dispendioso.

Em 2005 com o lançamento da primeira plataforma de sequenciação paralela massiva (a pirosequenciação) inaugurou-se uma nova era na análise de alto rendimento genómico (*high-throughput genomic analysis*) conhecida nos dias de hoje por sequenciação de nova geração (NGS, *next generation sequencing*). Nas últimas duas décadas têm-se assistido a um grande progresso nas técnicas de NGS permitindo a pesquisa de novas mutações em doenças mendelianas. A NGS veio revolucionar a investigação em genética, surgindo um crescente número de plataformas de NGS que permitem analisar rápida, eficazmente e a um custo cada vez menor toda a informação genética de um indivíduo. No entanto a análise bioestatística de todos os resultados obtidos continuam a ser um entrave ao avanço no conhecimento.

### **I.3.1 Sequenciação de Sanger**

Em 1977 foram publicados os primeiros dois artigos onde se descrevia uma técnica de sequenciação de DNA que iria revolucionar toda a investigação em genómica. Allan Maxam e Walter Gilbert descreveram um método que utiliza um processo de degradação química para cortar o DNA em pontos específicos, produzindo fragmentos de DNA de diversos tamanhos que eram posteriormente separados em eletroforese em

gel (Maxam and Gilbert 1977). Em alternativa a esta abordagem, Frederick Sanger e seus colegas descreveram o método de terminação de cadeias. Este método envolve a síntese de novo de uma série de cadeias simples de DNA, usando como molde a de DNA que se quer sequenciar. As cadeias sintetizadas são terminadas prematuramente nos vários tamanhos possíveis. A síntese começa sempre num ponto definido por um oligonucleótido iniciador (*primer*) e termina por incorporação de nucleótidos terminadores derivados dos dideoxi dos nucleótidos normais que não possuem um grupo hidroxilo na posição 3' da desoxirribose e impedem as ligações fosfodiéstericas do DNA (Sanger *et al.* 1977). Foram efetuados vários melhoramentos da técnica de Sanger o que levou a sua comercialização e ampla divulgação na comunidade científica, quer na investigação básica como no diagnóstico. A sequenciação de Sanger foi utilizada no projeto de sequenciação do genoma humano que durou 13 anos (Collins *et al.* 2003) (Frazier *et al.* 2003) (Human Genome Sequencing Consortium International 2004). Passadas mais de 3 décadas a sequenciação de Sanger continua a ser uma das metodologias mais utilizadas em investigação, mas principalmente em diagnóstico molecular, sendo o método preferencial para confirmação dos resultados obtidos pelas NGS. Com o avanço tecnológico hoje em dia existem sequenciadores para 96 ou 384 eletroforeses simultâneas, assim como o tamanho do fragmento a sequenciar e a especificidade da reação aumentaram grandemente e o custo diminuiu (Voelkerding *et al.* 2009).

### **I.3.2 Next Generation Sequencing**

A NGS em contraste à sequenciação de Sanger é uma técnica onde é possível sequenciar em paralelo vários fragmentos de DNA em oposição a um único possível na sequenciação de Sanger. Estes métodos são mais rápidos que a sequenciação por terminação da cadeia. A primeira plataforma de NGS a surgir foi 454 *life Sciences* (Roche®) que é uma técnica de pirosequenciação (Margulies *et al.* 2005; Wheeler *et al.* 2008).

#### **I.3.2.1 Pirosequenciação**

A pirosequenciação permite a sequenciação de fragmentos de 150bp em cada leitura, uma vez que podem ser feitas milhares de leituras em paralelo, o total de DNA

sequeenciado pode chegar a 1000Mb por corrida. No processo de pirosequenciação o DNA alvo é isolado ou obtido por PCR ou clonagem, e fragmentado em moléculas com 300bp a 800bp; Os fragmentos são tratados de modo a lhes serem adicionados dois adaptadores diferentes A e B às extremidades 3' e 5', respetivamente; os fragmentos são desnaturados e ligados a esferas de agarose que têm uma pequena sequência complementar a um dos adaptadores. Cada esfera deve estar apenas associada a um fragmento, o que se consegue através de sucessivas diluições. Adicionam-se ao meio reagentes de PCR e óleo, de modo a que se formem micelas que incluem uma esfera de agarose e reagentes. Cada uma destas micelas funciona como um microrreactor onde ocorre um PCR de emulsão. Após vários ciclos de PCR cada microrreactor vai conter, ligadas à esfera, milhares de cópias de um único fragmento. O *primer* utilizado é complementar ao adaptador que não está a ser utilizado para a ligação à esfera. Concluída a reação de amplificação remove-se o óleo e a solução é colocada numa placa que contém pequenos compartimentos onde apenas uma esfera pode ser inserida. Em cada compartimento vai ocorrer sequenciação do fragmento amplificado. A placa tem milhares de compartimentos, podendo sequenciar-se milhares de fragmentos em paralelo. A identificação dos nucleótidos é feita por quimioluminescência em simultâneo com a sua polimerização na reação de sequenciação. Sempre que é adicionado um nucleótido a uma cadeia em crescimento é libertado um pirofosfato (PPi) que reage com APS para originar ATP e sulfato (enzima ATP-sulfurilase). O ATP reage com a luciferina (enzima luciferase adicionada ao meio) e oxigénio numa reação que leva à emissão de luz. Em cada compartimento é adicionado um nucleótido de cada vez (os 4 dNTP's são adicionados sempre na mesma ordem) e verifica-se se há emissão de luz ou não. Antes da adição do nucleótido seguinte é necessário lavar ou inativar os dNTP's que ficaram em solução. Quando na sequência surge um nucleótido repetido (ex: GGGG) o sinal captado é superior ao esperado para um só nucleótido, sendo interpretado corretamente como uma repetição. Como consequência da pirosequenciação, a mesma região é sequenciada tantas vezes que, mesmo que ocorram erros de leitura ou de reação, a maioria das sequências estará correta, levando a um resultado global positivo e menos falível que a técnica clássica de terminação da cadeia (Shendure and Ji 2008).

### I.3.2.2 Whole-exome-sequencing

Estima-se que a maioria das mutações que originam doenças monogénicas se localiza nas regiões codificantes, os exões, ou em locais de *splicing* junto aos exões. Sendo assim, sequenciar apenas os exões é praticamente suficiente para que sejam detetadas mutações que causam este tipo de doença. É utilizado o termo exoma ou *whole-exome-sequencing* (WES) para esta técnica (Kuhlenbäumer *et al.* 2011). Esta é uma das técnicas mais utilizadas nos últimos anos e que demonstrou ser uma ferramenta poderosa para a deteção de mutações em doenças mendelianas onde a etiologia da doença era desconhecida até então (Majewski *et al.* 2011). Nos últimos anos diversos artigos têm sido publicados com estudos de investigação utilizando esta técnica para a deteção de doenças mendelianas com grande resultado (Kuhlenbäumer *et al.* 2011).

O método WES encontra-se bem estabelecido e baseia-se na sequenciação de pequenos fragmentos exónicos da molécula original, clonados em vetores, e na sobreposição de regiões comuns das diferentes sequências obtidas de modo a obter a sequência completa. Existem até ao momento 3 plataformas utilizadas para a captura e sequenciação do exoma: Roche/NimbleGen® (SeqCap EZ), Agilent Technologies® (*Sure-Select*) e Illumina® (*TruSeq Exome Enrichment*) (Erlich *et al.* 2008; Mardis 2008; Quail *et al.* 2008; Teer and Mullikin 2010; Ku *et al.* 2011). Todas estas plataformas, com algumas variações, seguem a seguinte estratégia: O DNA é fragmentado sendo depois adicionados adaptadores a ambas as extremidades dos fragmentos de DNA. Os fragmentos de DNA são amplificados por PCR através de *primers* biotinilados específicos para os exões. Os fragmentos de DNA resultantes são enriquecidos por hibridação com as sondas específicas, esta hibridação tem uma duração variável consoante a plataforma em questão, mas usualmente é realizada durante um período de 72 horas. Após esta hibridação utiliza-se pérolas de estreptavidina magnéticas para capturar o DNA que se encontra ligado especificamente às sondas sendo depois eluídas. O DNA capturado é depois novamente amplificado. Este enriquecimento é medido usando PCR em tempo real que avalia se o enriquecimento foi bem-sucedido e se está apto para passar ao passo seguinte, a sequenciação destes fragmentos. O objetivo é obter-se 20 a 30 vezes por base (chamado de profundidade de sequenciação) para que se considere que houve uma boa cobertura de todo o exoma. Usualmente a plataforma mais utilizada para esta sequenciação é o sequenciador Illumina® (Quail *et al.* 2008; Teer and Mullikin 2010).

### I.3.2.3 Whole-genome-sequencing

A sequenciação de genomas (WGS, *whole-genome-sequencing*) tal como o WES baseia-se na sequenciação de pequenos fragmentos da molécula original. Existem duas estratégias fundamentais para a sequenciação de genomas: A) *Whole genome shotgun sequencing* onde todo o genoma é fragmentado aleatoriamente em pequenos frações que são sequenciadas separadamente (Ng and Kirkness 2010). As sequências obtidas são analisadas na procura de regiões idênticas e o genoma é reconstruído pela sobreposição dessas regiões. Em genomas que contenham várias regiões idênticas é possível que a sequenciação não seja feita corretamente, sobrepondo-se sequências idênticas ou eliminando-se algumas regiões. Este tipo de sequenciação não é assim indicada para genomas grandes, como os de eucariotas; B) *Clone-by-clone shotgun sequencing*: cada cromossoma é tratado individualmente e fragmentado em várias frações da ordem dos megabits (Mb) (Pareek *et al.* 2011). Determina-se previamente a ordem desses fragmentos no genoma e só depois se aplica a técnica *shotgun* a cada fragmento. Uma vez sequenciados os fragmentos de maior dimensão a sequência do cromossoma é facilmente determinada. Embora mais fiável, este método é mais moroso e dispendioso que o anterior (Pareek *et al.* 2011).

A NGS trouxe uma nova abordagem na pesquisa de mutações em doenças mendelianas, no entanto e a apesar da eficiência e rapidez da técnica a quantidade de dados gerada por cada WES ou WGS é bastante elevada, variando consoante a cobertura analisada, plataforma de sequenciação e programa bioinformático utilizado para verificar as alterações. Normalmente são identificadas entre 20000 a 50000 alterações por WES (Zhang *et al.* 2011). De modo a se evitar um elevado número de falsos-positivos, todas as alterações passam vários filtros de controlo de qualidade. Após este filtro pode-se optar por outros na tentativa de identificar a alteração causadora de doença (Pop and Salzberg 2008; Yang *et al.* 2009; Gilissen *et al.* 2012).

No futuro, o WES e o WGS serão os métodos preferenciais para investigação mas sobretudo para o diagnóstico molecular, que poderá ser mais exaustivo.

## **I.4 Objetivos**

As doenças cardiovasculares (DCV) são a principal causa de morte na maioria dos países ocidentais. As DCVs têm etiologia e patologia complexas, resultantes da

combinação de fatores ambientais (tabagismo, sedentarismo, dieta inadequada, *stress*), biológicos (colesterol, triglicéridos, etc) e genéticos. No entanto a maioria destes fatores de risco é modificável. É estimado que 5-10% dos doentes que sofrem um enfarte do miocárdio antes dos 55 anos de idade possam ter FH (Marks *et al.* 2000; Neil *et al.* 2003). Doentes com FH corretamente identificados e tratados em idade jovem apresentam uma esperança de vida semelhante a qualquer outro indivíduo da população em geral; doentes com FH não tratados adequadamente podem ver a sua esperança de vida diminuída em 20 a 30 anos (Neil *et al.* 2003) o que demonstra a importância do diagnóstico precoce e terapêutica adequada desta patologia para a prevenção cardiovascular. É importante distinguir uma hipercolesterolemia hereditária de uma hipercolesterolemia não hereditária (poligénica/ambiental), uma vez que a FH se associa a um elevado risco de doença coronária prematura, fatal e não fatal, pois o indivíduo está exposto a níveis elevados de colesterol plasmático desde o nascimento. Sabendo que doentes com FH têm um risco cem vezes superior de desenvolver DCV prematura é importante identificar a causa genética destes indivíduos (Simon Broome Group 1991).

Com o intuito de contribuir para a diminuição da mortalidade cardiovascular e para a prevenção da doença, o presente trabalho de doutoramento teve como objetivo elucidar a causa genética da hipercolesterolemia em doentes com diagnóstico clínico de FH nos quais foi identificada uma alteração cuja patogenicidade era desconhecida ou nos quais não foi possível identificar uma mutação num dos 3 genes associados a Hipercolesterolemia Familiar. Para atingir este objetivo foram delineados os seguintes objetivos:

- a) Reorganização do estudo molecular incluindo o estudo de grandes rearranjos no *LDLR* e a caracterização funcional por estudos *in silico* e *in vitro* de alterações encontradas nos genes associados a FH no âmbito do EPHF mas cuja patogenicidade era desconhecida.
- b) Análise dos genes já associados à FH (genes *LDLR* e *APOB*) através da utilização de técnicas de *Next generation sequencing* com a perspetiva de excluir a possibilidade da existência de mutações no *LDLR* não identificadas pelas técnicas de biologia molecular utilizadas em protocolos de rotina e com o objetivo de estender o diagnóstico molecular da FH a todo o gene da *APOB*. Caracterização funcional de novas alterações no gene *APOB*.

- c) Avaliação bioquímica de casos índice do EPFH com e sem mutação (1999-2012).
- d) Sequenciação de todo o exoma de 5 doentes clínico de FH e hipercolesterolemia severa de forma a identificar novos genes associados a FH.







## CAPÍTULO II

---

### MATERIAL & MÉTODOS



## II MATERIAL E MÉTODOS

### II.1 Recrutamento de doentes

Todos os doentes estudados foram recrutados para o Estudo Português de Hipercolesterolemia Familiar (EPHF) entre 1999 e 2012. A inclusão de casos índice com diagnóstico clínico de FH neste estudo requer que sejam verificados os critérios abaixo mencionados para FH confirmada ou possível, adaptados dos critérios de “*Simon Broome Heart Research Trust*” (Group 1991). Alguns casos índice foram no entanto incluídos no EPHF mesmo sem cumprirem todos os critérios de admissão ao estudo, uma vez que o diagnóstico clínico é feito por um médico e sempre que o médico recomendava o estudo o caso índice era incluído no estudo.

Tabela II.1 – Critérios de inclusão no EPHF adaptados de “*Simon Broome Heart Research Trust*” para o diagnóstico de hipercolesterolemia familiar heterozigótica.

Critério	
<b>Determinações bioquímicas</b>	<b>A.</b> CT>290 mg/dL (>7.5 mmol/L) ou c-LDL>190 mg/dL (>4.9 mmol/L) num <u>adulto</u> ou CT>260 mg/dL (>4.0 mmol/L) ou c-LDL>155 mg/dL (>4.0 mmol/L) numa <b>criança menor de 16 anos</b>
<b>Sinais Físicos</b>	<b>B.</b> Xantomas tendinosos no caso índice ou em algum familiar em primeiro <sup>*</sup> ou segundo grau <sup>**</sup>
<b>História familiar</b>	<b>C.</b> Enfarte agudo do miocárdio em familiar em primeiro <sup>*</sup> ou segundo grau <sup>**</sup> com menos de 50 anos ou em familiar em primeiro grau <sup>*</sup> com menos de 60 anos
	<b>D.</b> CT>290 mg/dL (>7.5 mmol/L) num familiar em primeiro <sup>*</sup> ou segundo grau <sup>**</sup>
<b>Estudo molecular</b>	<b>E.</b> Evidência genética de mutação no gene <i>LDLR</i> ou no gene <i>APOB</i>
<b>Diagnóstico:</b>	<b>HeFH confirmada (<i>Definite HeFH</i>) → A+B ou E</b> <b>HeFH provável (<i>Probable HeFH</i>) → A+C ou A+D</b>

<sup>\*</sup>, familiar em primeiro grau: pais, filhos, irmãos; <sup>\*\*</sup>, familiar em segundo grau: avós, netos, sobrinhos, sobrinhas, tios, tias, meios-irmãos. Abreviaturas: c-LDL, colesterol associado às lipoproteínas de baixa densidade; CT, colesterol total; HeFH, hipercolesterolemia familiar heterozigótica.

Para o grupo pediátrico o valor de IMC foi calculado para género e idade com base nas curvas de crescimento do *Centre for Disease Control* (Kuczmarski *et al.* 2002)

recomendado pela Direcção Geral da Saúde. A pré-hipertensão e a hipertensão (sistólica e diastólica) foram definidas com base na idade e género, sempre que se verificaram os percentis 95 e 99, respetivamente, com base nas tabelas do “*The Fourth Report on the Diagnosis, Evaluation, and Treatment of High Blood Pressure in Children and Adolescents*” (National High Blood Pressure Education Program Working Group on High Blood Pressure in Children and Adolescents. 2004) recomendadas pela Direcção Geral da Saúde.

Para o grupo adultos o valor de IMC foi calculado tendo como base as recomendações da OMS. A hipertensão foi definida com base num valor superior ou igual a 140/90 (sistólica / diastólica), conforme recomendado pela ESC (Mancia *et al.* 2007)

## **II.2 Determinações bioquímicas**

As determinações dos parâmetros lipídicos, nomeadamente colesterol total, colesterol LDL, colesterol HDL, triglicéridos, apolipoproteína AI, apolipoproteína B e lipoproteína (a) de todos os indivíduos referenciados ao EPFH foram realizadas entre 1999 e 2008 num autoanalisador Hitachi 911 (Boehringer Mannheim, Roche) e entre 2009-2012 num autoanalisador Cobas Integra 400 Plus (Roche) através de métodos enzimáticos e colorimétricos. Estas determinações foram realizadas por técnicos do laboratório de Química Clínica da Unidade Laboratorial Integrada (ULI) do INSA.

A quantificação de apolipoproteína E, apolipoproteína A-II, apolipoproteína C-II, apolipoproteína C-III e do colesterol associado às sdLDL (*small dense low density lipoprotein*) foram realizadas no equipamento Daytona (Randox) existente no laboratório de investigação, a todos os casos índice referenciados ao EPFH e a todos os familiares de casos índice identificados molecularmente.

## **II.3 Colheita e processamento de amostras**

Para cada caso índice e respetivo familiar foram colhidas amostras de sangue em tubo EDTA (10 mL para o caso índice e 5 mL para familiar) para extração de DNA e 7,5 mL de tubo gel para obtenção de soro para determinação dos parâmetros bioquímicos. Estas amostras foram enviadas por correio ou colhidas no INSA, juntamente com um questionário preenchido pelo médico assistente e uma declaração de consentimento informado e esclarecido assinada por cada caso índice (no caso de o caso índice ser um menor o consentimento informado é assinado pelos progenitores ou representante legal)

e por cada familiar. A cada amostra foi atribuído um número de identificação e todas as informações relativas ao caso índice e seus familiares encontram-se registradas numa base de dados confidencial que se encontra legalizada pela Comissão Nacional de Proteção de Dados (CNPd).

### **II.3.1 Extração de DNA**

O DNA genómico foi extraído a partir de leucócitos de uma amostra de 5 mL de sangue periférico, colhido em tubos contendo EDTA. Neste método de extração de DNA foi utilizado o *Wizard® Genomic DNA Purification Kit* (Promega) de acordo com as instruções do fabricante. O processo consiste em quatro etapas: lise dos eritrócitos, na qual se adiciona 9 mL da solução de lise de células (*Cell Lysis Solution*) à amostra de sangue total, num tubo de 15 mL. Esta mistura é homogeneizada por inversão e incubada 10 minutos à temperatura ambiente com agitação moderada. Após uma centrifugação a 2000 x g durante 10 minutos a 18°C (centrifuga 5810R, *Eppendorf*), o sobrenadante é removido por decantação e o *pellet* (que contém os leucócitos) é agitado no vortex até estar totalmente dissolvido. Em seguida, adiciona-se 3 mL de *Nuclei Lysis Solution* e agita-se 2 a 3 vezes por inversão, permitindo a lise dos leucócitos e dos seus núcleos. O passo seguinte consiste na precipitação das proteínas celulares, através da adição de 1 mL de *Protein Precipitation Solution*. Esta mistura é agitada, primeiro por inversão (4 vezes) e em seguida no vortex durante 20 segundos. Após centrifugação 2000 x g durante 10 minutos a 18°C forma-se um *pellet* de detritos celulares e proteínas, sendo o sobrenadante transferido para um novo tubo de 15 mL. O último passo consiste na precipitação do DNA, por meio da adição de Isopropanol (Merck) ao sobrenadante, agita-se a mistura lentamente por inversão até que sejam visíveis as fibrilhas de DNA. O DNA é removido diretamente com uma micropipeta e colocado num tubo de 2 mL que contém 1 mL de etanol a 70% (v/v) e a -20°C, para permitir a remoção de sais. Em seguida centrifuga-se a 14000 x g durante 10 minutos à temperatura ambiente (microcentrifuga 5415D, *Eppendorf*), remove-se o sobrenadante repete-se a lavagem com etanol. Depois de seco o DNA é re-hidratado em 200µL de água estéril bidestilada e colocado num agitador (Thermomixer compact, *Eppendorf*) durante 12 horas a 450 rpm. Após estar completamente dissolvido uma alíquota de DNA é conservada a 4°C e outra a -20°C.

### II.3.2 Extração de RNA

O RNA foi extraído a partir de células mononucleares de uma amostra de 6 mL de sangue periférico, colhido em tubos de CPT. Neste método de extração utilizou-se (RNeasy Mini Kit, Qiagen) de acordo com as indicações do fabricante. O processo consiste nas seguintes etapas: após a colheita do sangue periférico para o tubo CPT centrifuga-se os tubos, num espaço máximo de 2 horas a  $1578 \times g$  durante 30 minutos à temperatura ambiente (centrifuga 5810R *Eppendorf*). Ressuspende-se o sobrenadante por inversão e transfere-se para um tubo novo de 15 mL. Após a centrifugação a  $400 \times g$  durante 10 minutos a  $4^{\circ}\text{C}$ , procede-se à contagem das células numa câmara de Neubauer juntando-se 1  $\mu\text{L}$  da suspensão anterior e 9  $\mu\text{L}$  de um corante (Trypan Blue, Gibco), contando-se apenas as células de um quadrante. O número de células é dado pela fórmula: Número de células = número de células contada  $\times 10^5$ . O passo seguinte é a extração imediata do RNA adicionando-se 350  $\mu\text{L}$  (no caso de o número de células ser  $<5 \times 10^6$ ) ou 600  $\mu\text{L}$  (número de células  $5 \times 10^6 - 1 \times 10^7$ ) de tampão RLT com 1% de  $\beta$ -Mercaptoetanol (Merck) homogeneizando-se o lisado com uma agulha de 20G 5 a 10 vezes. Transfere-se o lisado para uma coluna roxa QIAshredder colocada num tubo coletor de 2 mL. Após centrifugação a 13000 rpm durante 2 minutos à temperatura ambiente (centrifuga 5415R, *Eppendorf*), adiciona-se o igual volume de etanol a 70% ao volume de lisado sendo homogeneizado 10 vezes. Do volume total transfere-se 700  $\mu\text{L}$  da amostra para uma coluna rosa do Kit RNeasy colocada num tubo coletor de 2 mL sendo seguidamente centrifugado a 11000 rpm durante 15 segundos à temperatura ambiente, desprezando-se o eluído no final da centrifugação. Repete-se este passo 2 vezes. No passo seguinte adiciona-se tampão RW1 (permite remover contaminantes, contendo guanidina) e centrifuga-se nas mesmas condições anteriores. Após se desprezar o eluído adiciona-se 80  $\mu\text{L}$  de uma mistura de DNase (permite eliminar o DNA presente na amostra) na membrana de sílica-gel da coluna RNeasy e incuba-se à temperatura ambiente durante 15 minutos. Adiciona-se novamente RW1 e centrifuga-se nas mesmas condições. No passo seguinte despreza-se o tubo coletor com o eluído e transfere-se a coluna para um novo tubo (o tampão RW1 remove a DNase enquanto o RNA permanece ligado à membrana) ao qual se adiciona 500  $\mu\text{L}$  tampão RPE (permite remover contaminantes) centrifugando-se nas mesmas condições. Desprezou-se o eluído e adiciona-se novamente tampão RPE. Após centrifugação a 10000 rpm durante 2 minutos, transfere-se a coluna para um novo tubo coletor e repetiu-se a centrifugação. Após se desprezar o tubo coletor com o eluído transferiu-se a coluna para um novo tubo



de 1,5 mL e adicionou-se 50 µL *RNase free water*. Após centrifugação a 10000 rpm durante 1 minuto a 4°C desprezou-se a coluna. O RNA extraído encontra-se no tubo, retirou-se 3 µL para quantificação sendo o restante conservado a -80°C até posterior utilização para síntese de cDNA (ácido Desoxirribonucleico complementar).

A quantificação do RNA foi feita no equipamento Nanodrop 1000 (Thermo Scientific)

## **II.4 Diagnóstico molecular da FH**

### **II.4.1 Reorganização do diagnóstico molecular**

Durante este trabalho dividiu-se as fases do diagnóstico molecular da FH em 5. A fase I foi implementada antes do início deste estudo, mas todas as restantes fases descritas foram implementadas durante este trabalho.

Fase I – No período em que reporta este trabalho (2008-2012), esta fase foi realizada por um assistente de investigação do grupo de investigação cardiovascular e compreendeu as seguintes etapas: extração de DNA de todos os casos índice e familiares referenciados ao EPHF. Pesquisa de mutações nos exões 26 e 29 do gene *APOB* por amplificação de fragmentos que contêm parte dos exões 26 e 29 através da metodologia de PCR (*Polymerase Chain Reaction*) seguido de sequenciação direta. Em seguida foi feita a pesquisa de mutações em todos os 18 exões e região do promotor do gene *LDLR* pela técnica PCR seguido de sequenciação direta.

Fase II – Nesta fase foi realizada a pesquisa de grandes rearranjos (deleções ou duplicação) no gene *LDLR* pela técnica de MLPA (*Multiplex Ligation-dependent Probe Amplification*) dos casos índice sem mutação detetada no passo anterior ou com mutação já identificada mas com um fenótipo agressivo.

Fase III – Pesquisa de mutações em todos os casos índice sem mutação encontrada nas fases I e II de 4 exões do gene *PCSK9* onde se encontram descritas mutações causadoras de FH, pela metodologia de PCR seguida de sequenciação direta. Nos casos índice com fenótipo agressivo foi realizado o estudo completo do gene *PCSK9*.

Fase IV – Estudos funcionais de RT-PCR para verificação da patogenicidade das alterações de *splicing* e promotor do gene *LDLR*. O estudo da patogenicidade das alterações de *splicing* e promotor foi feito por RT-PCR e seguidamente os fragmentos foram analisados por sequenciação direta.

Fase V – Estudos funcionais de possíveis mutações no gene *LDLR* e possíveis mutações no gene *APOB*. Para os ensaios funcionais no gene *LDLR* procedeu-se ao isolamento e cultura de linfócitos dos casos índice com alteração a estudar. Os estudos de expressão, ligação e internalização foram efetuado por citometria de fluxo.

#### **II.4.2 Amplificação de DNA genómico por PCR**

Os fragmentos que contêm parte dos exões 26 e 29 do gene *APOB*, o promotor, os 18 exões do gene *LDLR* e os 12 exões do gene *PCSK9* foram amplificados do DNA genómico através da metodologia de PCR. Os oligonucleótidos iniciadores (*primers*) utilizados, bem como as diferentes temperaturas de hibridação dos *primers* encontram-se na tabela em anexo (Anexo A.I e A.2).

Para amplificação dos fragmentos dos genes *APOB*, *LDLR* e *PCSK9* utilizou-se a reação de PCR com o seguinte protocolo: a cada tubo de reação adicionou-se 0.2mM de cada desoxinucleótidos (dATP, dCTP, dGTP, dTTP, 100mM, Bioline), 5 tampão  $\text{NH}_4$  (10x, Bioline), 1.5 mM de  $\text{Mg}^{2+}$  (50 mM, Bioline), 20  $\mu\text{mol}$  de cada *primer* (Invitrogen), 1.25 U de *BioTaq* polimerase (Bioline), 100-200 ng de DNA genómico e água estéril bidestilada para perfazer o volume final de 50 $\mu\text{L}$ . os tubos foram colocados no termociclador (T300 e *TGradient*, Biometra) e a reação de PCR ocorreu nas seguintes condições: desnaturação inicial a 95°C durante 3 minutos, 35 ciclos de desnaturação a 94°C durante 45 segundos seguido de hibridação a 57°C-63°C, dependendo do exão a ser amplificado (Anexo A.I e A.2), durante 30 segundos e extensão a 72°C durante 1 minuto, extensão final a 72°C durante 15 minutos. Os produtos de amplificação foram guardados a 4°C até posterior utilização. A análise dos fragmentos foi realizada através de uma eletroforese em gel de agarose (ver II.4.3). Para a amplificação de alguns exões do gene *PCSK9* adicionou-se um adjuvante da PCR, o dimetilsulfóxido (DMSO).

#### **II.4.3 Análise dos fragmentos de DNA por eletroforese em gel de agarose**

Preparou-se um gel de agarose (Citomed) em 100 mL de tampão Tris/Borato/EDTA (TBE) 1 x (TBE 10 x ultrapure, pH 8.4, 1.0M Tris, 0.9M Ácido Bórico, 0.01M EDTA, Invitrogen), de modo a obter-se a concentração desejada (p/v). A concentração do gel varia consoante o fragmento que se pretende analisar, variando entre 0.8 % para a

avaliação qualitativa do DNA, 1.5% para a visualização dos fragmentos do produto PCR ou 2-3% para visualizar os fragmentos do produto RT-PCR. A solução foi aquecida no micro-ondas durante cerca de 2 minutos e após ser arrefecida adicionou-se syber safe ( $\mu\text{g/mL}$ ). Deixou-se solidificar cerca de uma hora. As amostras (até  $10\mu\text{L}$ ) foram aplicadas no gel após misturadas com  $3\mu\text{L}$  de solução de deposição (Anexo xxx). No gel aplicou-se ainda um marcador de peso molecular consoante o fragmento que se pretendeu analisar (1Kb DNA Ladder, Boehringer Mannheim ou PUC 18, Invitrogen). Efetuou-se a migração electroforética, em tampão TBE a uma voltagem (fonte de voltagem Bio-Rad Power Pac 3000) e período de tempo necessários para a obtenção de uma boa resolução das bandas do marcador de peso molecular (por ex. a migração de um gel a 1.5% é realizada a 70 volts durante 40 minutos). O registo do gel é feito em fotografia (Nikon E8700) usando o sistema de imagem safe imager (Invitrogen).

#### II.4.4 Sequenciação automática - método de Sanger

Antes de se realizar a reação de sequenciação, o excesso de *primers* e de dNTPs resultantes da reação de PCR devem ser eliminados através de uma digestão enzimática utilizando a combinação de duas enzimas hidrolíticas, a Exonuclease I e a *Shrimp Alkaline Phosphatase* (SAP), combinadas num produto comercial designado ExoSAP-IT (Amersham Pharmacia Biotech).

A um tubo de reação adicionou-se  $5\mu\text{L}$  de produto de PCR e  $2\mu\text{L}$  de ExoSAP-IT (20U de Exonuclease I e 4 U SAP), que foi colocado num termociclador (T3000, Biometra) nas seguintes condições: incubação a  $37^\circ\text{C}$  durante 15 minutos e inativação das enzimas a  $80^\circ\text{C}$  durante 15 minutos. Os produtos são guardados a  $4^\circ\text{C}$  até posterior utilização ou usado de seguida.

A reação de sequenciação foi preparada de acordo com as instruções do *Big Dye® Terminator Cycle Sequencing Ready Reaction Kit* (versão 1 ou 2, Applied Biosystems). A cada tubo de reação adicionou-se  $2\mu\text{L}$  de *Terminator Ready Reaction mix*,  $1\mu\text{L}$  do produto de PCR purificado e água estéril desionizada para perfazer o volume final de  $10\mu\text{L}$ . Os tubos foram colocados no termociclador (T3000, Biometra) e a reação de sequenciação ocorreu nas seguintes condições: desnaturação inicial a  $96^\circ\text{C}$  durante 30 segundos, 25 ciclos de desnaturação a  $96^\circ\text{C}$  durante 10 segundos seguido de hibridação a  $58^\circ\text{C}$  durante 5 segundos e extensão a  $60^\circ\text{C}$  durante 4 minutos. Os produtos são

guardados a 4°C até posterior utilização. Este produto foi depois entregue à Unidade de Tecnologia e Inovação (INSA) para ser sequenciado (3100 Genetic Analyser, Applied Biosystems) e as sequências obtidas foram analisadas com o programa *Staden Package*.

#### **II.4.5 Análise de grandes rearranjos no gene *LDLR* por *multiplex ligation-dependent probe amplification* (MLPA)**

Esta metodologia permite realizar a pesquisa de grandes rearranjos (deleções ou duplicações de um ou mais exões). Este método baseia-se na detecção de alterações no número de cópias numa única reação de PCR, utilizando um único par de *primers*. Neste método utilizou-se o kit *SALSA® MLPA® kit* (MRC-Holland), de acordo com as recomendações do fabricante. Nesta técnica não são as sequências alvo que são amplificadas mas as sondas de MLPA que hibridam com a sequência alvo. O kit *MLPA® Salsa P062-C2 LDLR* contém 33 sondas com produtos de amplificação entre 136 e 445 nucleótidos, 20 das quais para o gene *LDLR*. O kit possui também 11 sondas de referência, fragmentos controlo que originam pequenos produtos que são posteriormente usados para aferir sobre a qualidade do DNA da amostra, bem como garantir que a corrida e a reação de ligação sejam realizadas com sucesso, 2 sondas que emparelham imediatamente antes e depois do gene *LDLR*. Cada sonda MLPA consiste num par de *primers* (sondas de hibridação) marcados por fluorescência que reconhecem sequências-alvo adjacentes no DNA. Cada um dos *primers* contém uma das sequências reconhecidas pelo *primer* de PCR.

O processo consiste em quatro etapas: iniciou-se com a desnaturação do DNA, colocou-se 2.5 µL de DNA da amostra a analisar e 2.5 µL de TE num tubo de reação durante 5 minutos a 98°C no termociclador. Após a desnaturação, adicionou-se, no termociclador, a cada tubo 1.5 µL de tampão MLPA e 1.5 µL de uma solução que contém as sondas, seguindo-se um pequeno passo de desnaturação durante 1 minuto a 95°C e posteriormente um tempo de hibridação a 60°C durante 16 a 20 horas. Após a hibridação das sondas seguiu-se a reação de ligação, adicionou-se a cada tubo 32 µL do produto da reação de ligação, composto por 25 µL de água estéril bidestilada, 3 µL de tampão ligase A, 3 µL de tampão ligase B e 1 µL de enzima Ligase-65. A ligação ocorreu a 54°C durante 15 minutos seguindo-se a inativação da enzima a 98°C durante 5 minutos. O último passo é a reação de amplificação por PCR de cada sonda, neste passo foi necessário preparar uma solução com 7.5 µL de água estéril bidestilada, 2 µL

SALSA PCR *primers* e 0.5 µL de SALSA polimerase e adicionou-se 10 µL desta solução a cada tubo e a reação de PCR ocorreu nas seguintes condições: 35 ciclos de desnaturação a 95°C durante 30 segundos, seguido de hibridação dos *primers* a 60°C durante 30 segundos e extensão a 72°C durante 1 minuto, e extensão final a 72°C durante 20 minutos. Em cada reação foram colocadas amostras controlo conhecidas sem rearranjos no gene *LDLR*. O produto final foi entregue à Unidade de Tecnologia e Inovação (INSA) para ser analisado no sequenciador automático (ABI PRISMTM 3100 Genetic Analyser, Applied Biosystems). Os produtos de amplificação foram analisados pelo *software* de análise GeneScanner, versão 3.1.2, sendo o tamanho dos picos medido pelo *software* Genotyper, versão 2.5 (Applied Biosystems) e o resultado exportado para um documento em formato excel para posterior análise. A altura dos picos da sequência-alvo foi normalizada e avaliada quanto a perda (razão de 0.5) ou ganho de exão (razão de 1.5).

## **II.5 Next generation sequencing**

### **II.5.1 Pirosequenciação**

Foram estudados 65 doentes referenciados ao EPHF com diagnóstico clínico de FH e sem alteração nos genes *LDLR* e *PCSK9* e nos fragmentos do exão 26 e 29 do gene *APOB* estudados por rotina no nosso laboratório.

A pirosequenciação foi realizada pela companhia Biocant (Cantanhede, PT) no equipamento 454 Life Science (Roche®). Todos os exões do *LDLR* e *APOB* foram amplificados num pool de 20, 20 e 25 amostras com *custom design primers* e depois sequenciado. Foram obtidos para o gene *APOB* um total de 227688 *reads de* nucleótidos correspondendo a uma cobertura média de 35X/fragmentos/indivíduo. As sequências obtidas por pirosequenciação foram alinhadas com base em sequências consensos (*LDLR* - NC\_000019.9 and *APOB* - NC\_000002.11) utilizando o *software Amplicon Variant Analyzer* (AVA, Roche). Os resultados obtidos foram apresentados na forma de estimativa de alelos. A estimativa 1 alelo/alteração foi interpretado como pelo menos um alelo tem essa alteração. As alterações detetadas por pirosequenciação foram resequenciadas por sequenciação de Sanger para confirmação das alterações encontradas

### II.5.2 Sequenciação de exoma

A sequenciação do exoma foi efetuada pela empresa BaseClear B.V. (Netherland), os principais passos foram: o DNA foi fragmentado, o tamanho para amplificação foi selecionado e as extremidades das moléculas cortadas. Os adaptadores Illumina foram adicionados a ambas as extremidades dos fragmentos de DNA e as moléculas foram amplificadas utilizando as regiões adaptadoras nos locais de iniciação. Os fragmentos de DNA resultantes foram hibridados com sondas EZ SeqCap fornecidas pela Nimblegen. A hibridação decorreu durante 72 horas. Após a hibridação das sondas os fragmentos obtidos foram lavados e apenas os fragmentos ligados as sondas EZ SeqCap foram eluídos. O DNA eluído foi posteriormente amplificado por LM-PCR e o enriquecimento foi medido por qPCR. Os fragmentos foram posteriormente sequenciados no equipamento Illumina conforme descrição do fabricante Illumina®. Após a sequência os fragmentos obtidos passaram por um controlo de qualidade e apenas as sequências com boa qualidade foram consideradas para a análise. Os dados das sequências foram fornecidos pela empresa BaseClear no formato FastQ.

### II.6 Programas de análise *in silico*

Para a análise das alterações encontradas durante este trabalho utilizou-se diferentes programas de análise *in silico*. Para a análise de mutações pontuais, pequenas deleções ou inserções e alterações sinónimas utilizou-se os seguintes programas: *PolyPhen-2* (Adzhubei *et al.* 2010), *Sorting Tolerant From Intolerant (SIFT)* (Ng and Henikoff 2003), *Consensus Deleteriousness score of missense SNVs (Condel)* (González-Pérez and López-Bigas 2011), *Mutations taster* (Schwarz *et al.* 2010), *Grantham Score* (Grantham 1974) e *PhyloP* (Pollard *et al.* 2010). As alterações foram classificadas como possivelmente patogénica, provavelmente patogénica ou neutra, para o programa *PolyPhen-2*; como patogénica ou neutra para os programas SIFT e CONDEL; como causa doença ou polimorfismo para o *Mutations taster*. Para *Grantham Score*, a classificação é dada em termos numéricos: conservativa (0-50), conservativa moderada (51-100), radical moderado (101-150), ou radical ( $\geq 151$ ). O programa *PhyloP* dá uma classificação entre -14,1 e 6,4 e foi considerado como patogénico a partir de um *score* superior a 4,4.

Para a análise de mutações de *splicing* e sinónimas utilizou-se os seguintes programas: *Splice-Site Predictor (Splice Port)* (Dogan *et al.* 2007), *Neural Network Splice Site*

*Prediction Tool (NNSSP)* (Reese *et al.* 1997) e *neural network predictions of splice sites in human (NetGen2)* (Hebsgaard *et al.* 1996). Os programas utilizados produziram um resultado preditivo entre 0 a 100% (em que 0% afetava o correto processamento do pré-mRNA e 100% não afetava) para as regiões de *splice-site* (região dadora e recetora de *splicing*) onde se realizou a previsão tendo em conta a sequência de referência e a sequência com a alteração.

As mutações estão descritas de acordo com a nomenclatura recomendada pela *Human Genome Variation Society*, porposta por Taschner & Den Dunnen (Taschner and den Dunnen 2011). Para todos os genes estudados neste trabalho, a numeração foi baseada no cDNA considerando o nucleótido adenina como sendo o primeiro do codão de iniciação ATG. As sequências de referência utilizadas para os genes associados à FH encontram-se na tabela II.2, bem como as sequências das proteínas. Os genes do metabolismo lipídico considerados para a análise de dados do *exome sequencing* encontram-se no anexo A.4.

Tabela II.2 – Referências das sequências nucleóticas e proteicas dos genes e proteínas analisados neste trabalho.

Gene	NCBI	Uniprot
<i>LDLR</i>	NM_000527.4	P01130
<i>APOB</i>	NM_000384.2	P04114
<i>PCSK9</i>	NM_174936.3	Q8NBP7
<i>LDLRAP1</i>	NM_015627.2	Q5SW96

## II.7 Estudos funcionais

### II.7.1 Estudos de mutações *splicing* e promotor *LDLR*

#### II.7.1.1 Síntese de cDNA e amplificação do cDNA por PCR (Polymerase Chain Reaction)

A síntese de cDNA foi efetuada através da atividade da enzima Transcriptase Reversa, com recurso ao *kit* comercial TaqMan (Applied Biosystems) seguindo-se as recomendações do fabricante. A reação foi realizada para a obtenção de um volume final de 50 µL, onde se adicionou 1 µg de RNA, tampão RT 10x, 0,1 mM de cada

desoxinucleotídeos (dNTPs: dATP, dGTP, dTTP e dCTP, 2,5 mM cada), 0,275 mM MgCl<sub>2</sub> (25 mM), 0,00125 mM de Oligo d(T)<sub>16</sub> (50 µM, 5 nmoles), inibidor de RNase e 1,25 U/ µL de enzima Transcriptase Reversa. A reação foi efetuada num termociclador (T3000, Biometra), nas seguintes condições: 10 minutos a 25°C, 30 minutos a 48°C e 5 minutos a 95°C. As amostras de cDNA foram conservadas a -20°C até posterior utilização.

A amplificação do cDNA foi realizada da seguinte forma: a cada tubo de reação adicionou-se 0,2mM de cada desoxinucleótidos (dATP, dCTP, dGTP, dTTP, 100mM, Bioline), tampão NH<sub>4</sub> (10x, Invitrogen), 1,5 mM de Mg<sup>2+</sup> (50 mM, Invitrogen), 20 µmol de cada *primer* (Invitrogen), 0,8 U de Platinum® Taq DNA Polymerase, 100-200 ng de DNA genómico e água estéril bidestilada até perfazer um volume final de 50µL. Os tubos foram colocados no termociclador (T300 e TGradient, Biometra) e a reação de PCR ocorreu nas seguintes condições: desnaturação inicial a 94°C durante 2 minutos, 37 ciclos composto pelo passo de desnaturação a 95°C durante 45 segundos seguido de hibridação a 57°C-63°C, dependendo do exão a ser amplificado (Anexo A.3), durante 1 minuto e 45 segundos e extensão a 72°C durante 1 minuto, extensão final a 72°C durante 7 minutos. Os produtos de amplificação são guardados a 4°C até posterior utilização. A análise dos fragmentos foi realizada através de uma eletroforese em gel de agarose (ver II.4.4).

## **II.7.2 Estudos de mutações missense no *LDLR* e no gene *APOB***

### **II.7.2.1 Isolamento das partículas de LDL**

As partículas de LDL foram isoladas a partir de soro dos casos índice com a mutação na *APOB* e indivíduos normolipidémicos, após um jejum de 12 horas, a colheita foi realizada em tubos de gel apropriados para a obtenção de soro de acordo com o descrito anteriormente por Frostegard e seus colaboradores (Frostegard et al, 1990). Trinta minutos após a colheita foi efetuada uma centrifugação a 3500 rpm durante 15 minutos à temperatura ambiente (centrifuga 5810R, Eppendorf). O soro foi separado para um novo tubo e conservado a -80°C até posterior utilização. De modo a se obter as diferentes frações lipoproteicas foi realizada uma ultracentrifugação num rotor TST 41,14 a 35400 rpm, durante 19 horas à temperatura de 20°C (Centrikon T-2190, Kontron) das amostras de cada caso índice que permitiu separar as diferentes frações por densidade (Tabela II.3). De forma a se obter as frações lipoproteicas com uma boa



separação por densidade foi necessário ajustar, ao tubo de nitrato de celulose com o soro, Brometo de Potássio (KBr) (Merck) na razão 1,21 kg/L (por ex. uma amostra que tenha 4 mL de soro foi necessário ajustar com Brometo de potássio até se obter um peso de 4,84 gr/mL) e adicionou-se PBS (*phosphate-buffered saline*) a todos os tubos até apresentarem o mesmo peso. Após a ultracentrifugação recolheu-se para um tubo de 2 mL, com uma pipeta de Pasteur de vidro, a fração correspondente às partículas de LDL que foi conservada a 4°C no escuro até posterior utilização.

Tabela II.3 – Características das lipoproteínas humanas presentes no plasma.

Variável	Quilomicra	VLDL	IDL	LDL	HDL	Lp(a)
Densidade (g/mL)	<0.95	0.95-1.006	1.006-1.019	1.019-1.063	1.063-1.210	1.040-1.130
Peso Molecular (Da)	0.4-30x10 <sup>9</sup>	5-10 x10 <sup>6</sup>	3.9-4.8 x10 <sup>6</sup>	2.75 x10 <sup>6</sup>	1.8-3.6 x10 <sup>5</sup>	2.9-3.7 x10 <sup>6</sup>
Diâmetro (nm)	>70	25-75	22-24	19-23	4-10	25-30
Razão Lípido-proteína	99:1	90:10	85:15	80:20	50:50	75:25-64:34

VLDL, lipoproteínas de muita baixa densidade; IDL, lipoproteínas de densidade intermédia; LDL, lipoproteínas de baixa densidade; HDL, lipoproteínas de alta densidade; Lp(a), lipoproteína (a). (Adaptado de Rifai *et al.* 1999)

#### II.7.2.2 Purificação das partículas de LDL

Uma vez que a ultracentrifugação foi efetuada utilizando KBr foi necessário purificar-se as partículas de LDL de modo a não existirem contaminação deste composto. Este procedimento de purificação foi realizado em colunas apropriadas para diálise (PD-10 Desalting Columns, GE Healthcare). No primeiro passo adicionou-se, cerca de 5 vezes, água desionizada às colunas. Seguidamente adicionou-se Bicarbonato de Sódio (NaHCO<sub>3</sub> a 0,1M), também 5 vezes e por último adicionou-se todo o volume que foi recolhido de partículas de LDL. Após passagem na coluna, foram recolhidas as partículas de LDL para um novo tubo (as partículas de LDL apresentam um coloração amarela e apenas foi recolhido para o tubo as gotas que apresentavam essa coloração). Após esta recolha, quantificou-se as partículas de LDL de forma a conhecer-se a concentração existente por amostra. A concentração das proteínas foi determinada pelo método colorimétrico Bradford, a medição foi efetuada num espectrofotómetro, a 595nm (espectrofotómetro, 6300 Jenway) e calculou-se a curva padrão de BSA (*bovine*

*sérum albumin*) (Protein assay, Bio-Rad). Após a quantificação das partículas de LDL, colocou-se todas as amostras a uma concentração de 0,9 mg/ $\mu$ L.

#### II.7.2.3 Marcação das partículas de LDL

A marcação das partículas de LDL foi realizada com isotiocianato fluorescente (FITC) (*suitable for protein labeling* - Sigma-Aldrich) utilizado para ensaios em citometria de fluxo. Adicionou-se 7  $\mu$ L de marcador a 700  $\mu$ L de LDL a uma concentração de 0,9 mg/ $\mu$ L, incubou-se os tubos em agitação no escuro durante 2 horas. Após a incubação das partículas de LDL marcadas foram novamente purificadas em colunas de diálise. Os passos da purificação foram semelhantes aos anteriormente descritos, sendo a única alteração a adição de PBS em vez de  $\text{NaHCO}_3$ . Após a recolha das partículas de LDL marcadas os tubos foram conservados a 4°C no escuro até posterior utilização.

#### II.7.2.4 Isolamento e cultura de linfócitos

A separação de mononucleares foi realizada através de um gradiente originado por Ficoll numa centrífuga basculante de acordo com o descrito anteriormente por Etxebarria e seus colaboradores (Etxebarria *et al.* 2012). Colocou-se 4 mL de sangue periférico em cada tubo de 15 mL por cada 3 mL de ficoll. Colocou-se 3 mL de ficoll (Ficoll-Paque; GE Healthcare, Little Chalfont, UK) num tubo falcon de 15mL e adicionou-se 4 mL de sangue periférico muito lentamente para que não ocorresse mistura das fases. Após centrifugação a 400 x g durante 50 minutos a 18°C (centrifuga 5810R, Eppendorf) foi removido o anel branco para um novo tubo e lavou-se 2 vezes com a solução A (45mL tampão (NaCl filtrado) + 5 mL solução A (glucose 0,1%;  $\text{CaCl}_2$   $5 \times 10^{-5}\text{M}$ ,  $\text{MgCl}_2$   $9,8^{-4}\text{M}$ ; KCl 5,4mM, Tris 0,145M pH=0,5), nas seguintes condições de centrifugação, 150 x g durante 15 minutos a uma temperatura de 18°C. Eliminou-se o sobrenadante e adicionou-se 5 mL de meio *Dulbecco's Modified Eagle Medium* (DMEM) (Gibco, Invitrogen), com antibiótico e 10% *Lipoprotein Deficient Serum, human* (LPDS) para contagem de células. A contagem foi realizada num contador automático de células (Automated cell counter TC10, Bio-Rad). Os linfócitos, a uma concentração de  $2 \times 10^5$  células/mL, foram mantidos em crescimento em placas de 24 poços, a 37°C, 5%  $\text{CO}_2$  em meio DMEM com 10% de LPDS e os antibióticos penicilina e estreptomicina a 100 U/mL (Gibco, Invitrogen) durante 72 horas. De forma

a se obter uma maior ativação das células T, adicionou-se aos linfócitos *beads* Anti-CD3/CD28 I (Dynal Biotech, Oslo, Norway) a uma concentração de  $2 \times 10^5$  beads/mL.

O isolamento de linfócitos foi efetuado para amostras de indivíduos normolipidêmicos, casos-índice com alteração no *LDLR* caracterizada funcionalmente e para casos-índice com alterações no *LDLR* sem estudos funcionais.

#### II.7.2.5 Quantificação das partículas de LDL em linfócitos por citometria de fluxo

Após as 72 horas de crescimento dos linfócitos (descritas anteriormente) lavou-se os linfócitos 2 vezes com PBS contendo 1% *bovine serum albumin* (PBS-1%BSA), colocou-se o LDL marcado com FITC (partículas de LDL de indivíduos normais e de casos-índice com alterações no gene *APOB*) em cada poço e incubou-se durante 4 horas a 37°C para os ensaios de internalização e a 4°C para os ensaios de ligação LDL:LDLR. Após a incubação os linfócitos foram recolhidos para tubos de 1,5 mL e colocados num suporte eletromagnético para se retirar as *beads* de modo a não influenciar os resultados dos ensaios. Lavou-se novamente os linfócitos 2 vezes com PBS-1%BSA e fixou-se com 4% de paraformaldeído durante 10 minutos à temperatura ambiente. Seguidamente, efetuou-se nova lavagem com PBS-1%BSA, de forma a se eliminar o fixador. A intensidade de fluorescência foi medida por separação de células ativadas por fluorescência (FACS) num citómetro de fluxo (FACSCalibur Flow cytometer) de acordo com as indicações do fabricante. Para a análise de dados foi adquirida para cada amostra a fluorescência de 10,000 eventos e os dados foram expressos como percentagem dos valores máximos obtidos em células com partículas de LDL normais, sendo a média de fluorescência de células ativas bloqueadas, medidas no FACS segundo o *setting* frente contra a janela lateral de dispersão (*forward versus side-scatter window*). De forma a se medir a quantidade internalizada de partículas de LDL adicionou-se, às amostras para os ensaios de internalização, solução de azul tripano (0,2% concentração final) eliminando desta forma o sinal extracelular do complexo LDL:LDLR. Este passo permitiu remover a fluorescência extracelular através de um *quenching* e determinar a intensidade das partículas fluorescentes remanescentes no interior das células (que não foram afetadas pelo *quenching*). Os ensaios de ligação LDL:LDLR foram determinados por medição da intensidade de fluorescência associada à LDL marcada com FITC por FACS. Todas as medições foram realizadas em triplicado em, pelo menos, duas experiências independentes.

#### II.7.2.6 Quantificação da expressão do LDLR em linfócitos por citometria de fluxo

Após às 72h de incubação dos linfócitos, como descrito anteriormente em II.7.2.4 os linfócitos foram ressuspensos com PBS-1%BSA e fixou-se com 4% de paraformaldeído durante 10 minutos à temperatura ambiente. Seguidamente bloqueou-se as células com PBS com 10% de soro fetal de bovino (PBS-10%FBS) durante 1 hora a 4°C. Seguidamente lavou-se com PBS-1%BSA e incubou-se com um anticorpo primário de rato anti-LDLR (1:100; 2,5 mg/L; Progen Biotechnik GmbH) durante 1 hora a 4°C. Após a incubação lavou-se e vezes com PBS-1%BSA e incubou-se com um anticorpo secundário de cabra anti-IgG de rato conjugado com *Alexa Fluor 488* (1:100, Invitrogen, Molecular Probes, Carlsbad, USA). Para a análise de dados foi adquirida para cada amostra a fluorescência de 10,000 eventos e os dados foram expressos como percentagem dos valores máximos obtidos em células com partículas de LDL normais, sendo a média de fluorescência de células ativas bloqueadas, medidas no FACS segundo o *setting* frente contra a janela lateral de dispersão (*forward versus side-scatter window*). Todas as medições foram efetuadas em triplicado em, pelo menos, duas experiências independentes.

#### II.7.2.7 Quantificação das partículas de LDL em linha celular de hepatócitos (HepG2) por citometria de fluxo

As células HepG2 (liver hepatocellular cells) foram mantidas em crescimento em monocamada a 37°C, 5%CO<sub>2</sub> em meio DMEM com 10% de soro fetal de bovino (FBS Invitrogen), com os antibióticos penicilina e estreptomicina a 100 U/ml (Gibco, Invitrogen). As células foram semeadas em placas de 24 durante 24 horas antes dos ensaios de internalização. Para o ensaio de internalização colocou-se o LDL marcado com FITC em cada poço, para cada amostra a estudar, e incubou-se durante 4 horas a 37°C. Após a incubação lavou-se 2 vezes com PBS e fixou-se com 4% de paraformaldeído durante 10 minutos a temperatura ambiente. Seguidamente, efetuou-se nova lavagem com PBS, de forma a se eliminar o fixador. A quantificação foi efetuada em FACS tal como descrito anteriormente para a internalização nos linfócitos. Todas as medições foram efetuadas em triplicado em, pelo menos, duas experiências independentes.

#### II.7.2.8 Ensaio de proliferação das células U937

Células U937 (Human leukemic monocyte lymphoma cell line, ATCC® CRL-1593.2TM) foram mantidas em crescimento a 37°C, 5%CO<sub>2</sub> em meio RPMI-1640 (Roswell Park Memorial Institute, Sigma-Aldrich) com 10% de FBS e com os antibióticos ampicilina e estreptomicina a 100 U/ml (Gibco, Invitrogen) num frasco de cultura de células 75 a uma concentração de  $2 \times 10^5$  células/mL. Antes dos ensaios de proliferação as células foram semeadas numa placa de 96 poços durante 24 horas num meio RPMI com 10% de LPDS. O ensaio iniciou-se com a adição de 2µg/mL de partículas de LDL de cada amostra a estudar e incubou-se durante 48 horas a 37°C para que ocorresse o crescimento das células U937. A proliferação das células U937 foi determinada pelo método *CellTiter 96® AQueous Non-Radioactive Cell Proliferation Assay*. Este método baseou-se na taxa de redução do corante MTS [3-(4,5-dimetiltiazol-2-yl)-5-(3-carboximetoxifenil)-2-(4-sulfofenil)-2H-tetrazolium] na presença de um reagente de ligação de electrões PMS (metossulfato de fenazina, 5% (v/v) por desidrogenase mitocondrial das células metabólicas ativamente, produzindo um produto solúvel denominado de formazano que pode ser medido a 490 nm. A quantidade de formazano gerado foi diretamente proporcional ao número de células viáveis na cultura. Os resultados da proliferação das células U937 para as diferentes amostras de partículas de LDL dos doentes foi expressa em relação ao aumento do crescimento celular. Todas as medições foram efetuadas em triplicado em, pelo menos, duas experiências independentes.

Para todos os ensaios foram utilizados um controlo positivo para a mutação APOB3527 e um controlo negativo, ou seja, sem alterações no gene APOB.

#### II.8 Painel de controlos normolipidémicos

Foi construído um painel de controlos normolipidémicos com 96 amostras de indivíduos da população portuguesa com todos os parâmetros lipídicos dentro dos valores recomendados pela Sociedade Europeia de Cardiologia (Catapano *et al.* 2011a). Todas as alterações não descritas anteriormente e encontradas no decorrer deste trabalho nos diferentes genes foram pesquisadas, por PCR e sequenciação automática, neste painel de modo a se calcular a prevalências das mesmas.

### **II.9 Análise estatística**

A análise estatística foi realizada no programa *SPSS* (versão 17.0 para Windows; SPSS, Chicago, Illinois). A análise descritiva dos dados foi realizada tendo como intervalo de confiança 95%. A análise aos valores médios das variáveis quantitativas para amostras independentes foi realizada pelo teste t-student, enquanto o teste de Mann-Whitney ou Kruskal Wallis foi realizado para médias não-paramétricas. Para a análise das variáveis qualitativas aplicou-se o teste de independência do Qui-quadrado ( $\chi^2$ ).







## CAPÍTULO III

---

### RESULTADOS



## Parte 1 – Caracterização Molecular no âmbito do Estudo Português Hipercolesterolemia Familiar

Este trabalho está incluído nos seguintes artigos:

Bourbon M, **Alves AC**, *et al.* (2008) *Atherosclerosis* 196: 633–642

Medeiros AM, **Alves AC**, *et al.* (2010) *Atherosclerosis* 212:553–558

**Alves AC**, *et al.* (2013) *Hum. Mol. Genet* doi:10.1093/hmg/ddt573



## **Parte 1 – Caracterização Molecular no âmbito do EPHF.**

O Estudo Português de Hipercolesterolemia Familiar (EPHF) iniciou-se em 1999 e tem como objetivo a realização de um estudo epidemiológico para a determinação da prevalência e distribuição da Hipercolesterolemia Familiar em Portugal, tendo implementado o estudo molecular dos três genes associados à FH, o *LDLR*, a *APOB* e o *PCSK9*. Pretende também contribuir para a melhor compreensão da fisiopatologia da doença cardiovascular nestes indivíduos, com o intuito de melhorar o seu prognóstico (Bourbon and Rato 2006). Sabe-se que a identificação precoce da FH leva ao aconselhamento e tratamento adequado, minimizando o risco de eventos cardiovasculares e podendo mesmo levar à regressão da aterosclerose (Marks *et al.* 2003). O diagnóstico molecular da FH permite a correta identificação da patologia e fundamenta a aplicação de uma terapêutica farmacológica mais agressiva e/ou precoce, com a consequente redução do risco cardiovascular nos indivíduos afetados (NICE 2008).

Para melhor compreensão dos resultados dividiu-se o EPHF em duas fases: Fase A entre os anos de 1999 e 2007 e fase B de 2008 a 2012, período no qual decorreu o presente trabalho de doutoramento. Os resultados de todo o trabalho experimental serão descritos em pormenor nas referentes alíneas.

### **III.1.1 Período A**

Durante o período A foram identificados 137 casos-índice (CI) com pelo menos uma alteração num dos 2 genes associados à FH, o *LDLR* e *APOB*: 125 CI apresentavam uma alteração em heterozigotia no gene *LDLR* e 7 CI tinham uma alteração no gene *APOB*. Para além dos 125 CI identificados com uma mutação em heterozigotia no gene *LDLR* foram também identificados 2 CI homozigóticos e 3 CI heterozigóticos compostos para o gene *LDLR*. No entanto, 119 CI não apresentavam nenhuma alteração num dos 2 genes estudados.

### III.1.2 Período B

Entre os anos de 2008 e 2012, período em que decorreu este trabalho, foram estudados 388 CI, tendo sido identificadas alterações em heterozigotia nos três genes associados à FH: 136 CI no gene *LDLR*; 5 CI no gene *PCSK9* e 18 CI no gene *APOB*. Foram também identificados 1 CI homozigótico e 2 CI heterozigóticos compostos para o gene *LDLR*. Para além do estudo dos CI recebidos entre 2008 e 2012 foi também realizado no período B o estudo molecular para confirmação da suspeita clínica em alguns familiares de CI identificados no período A.

#### III.1.2.1 Pesquisa de mutações nos genes *LDLR* e *APOB* (fragmentos do exão 26 e 29) no âmbito do EPHF

A pesquisa de mutações dos genes *LDLR* e *APOB* aos CI referenciados ao EPHF foi realizada por um assistente de investigação do grupo de investigação cardiovascular e enquadra-se na fase I do diagnóstico molecular do EPHF. Sempre que foi identificada uma alteração num CI foi realizado o estudo molecular aos familiares disponíveis.

#### III.1.2.2 Estudo de grandes rearranjos pela técnica de MLPA

Para o estudo de grandes rearranjos foram analisados por MLPA 312 CI sem mutação detetada nos genes estudados na fase I do diagnóstico molecular do EPH (CI recebidos entre 1999-2012), bem como CI com uma mutação já identificada mas com um fenótipo agressivo. Foram identificados 6 grandes rearranjos, diferentes, em 19 CI (Tabela III. 1.1). A deleção do promotor ao exão 2 e do exão 8 ao exão 12 é o grande rearranjo mais frequente e uma das mutações mais comuns do EPHF com uma frequência de aproximadamente de 6%.

#### III.1.2.3 Pesquisa de mutações no gene *PCSK9*

Os CI sem mutação detetada nas fases I e II de diagnóstico molecular no âmbito do EPHF entre 1999-2012, foram analisados para o gene *PCSK9*. No total foram estudados 272 CI, uma vez que em 82 CI foi realizado o estudo completo do gene *PCSK9* e em 190 CI, foram pesquisadas apenas alterações nos exões 2, 4, 7 e 9, exões onde foram descritas mutações funcionais que originam um fenótipo de hipercolesterolemia. Como descrito por Abifadel *et al* (Abifadel *et al.* 2009), o gene *PCSK9* é um gene polimórfico

cujas alterações podem apresentar fenótipos distintos. De forma a caracterizar molecularmente os CI em estudo, foram registadas todas as alterações associadas quer ao fenótipo de hipercolesterolemia como ao de hipocolesterolemia (Tabela III.1.2).

Tabela III.1.1 – Grandes rearranjos identificados nos CI do EPHF

Localização no cDNA	Previsão na proteína	Co segregação‡	Referência
c.-30-?_190+?del + 1061-?_1845+?del	[Pr_Ex2del] + [Ex8_12del]	39/39; 0/14	(Bourbon <i>et al.</i> 2008)
c.68-?_313+?del	Ex2_Ex3del	5/5;0/0	(Ma <i>et al.</i> 1989; Yamakawa <i>et al.</i> 1989; Hobbs <i>et al.</i> 1992; Medeiros <i>et al.</i> 2010)
1061-?_1845+?del	Ex8_Ex12del	1/1;0/0	(Bourbon <i>et al.</i> 2008)
1587-?_1845+?del	Ex11_Ex12del	1/1;0/1	(Bertolini <i>et al.</i> 1995)
1846-?_2311+?dup	Ex13_Ex15Dup	4/4;0/0	(Lelli <i>et al.</i> 1991)
c.2312-?_*220+?del	Ex16_Ex18del	2/2;0/0	(Medeiros <i>et al.</i> 2010)

‡ Co-segregação: número de indivíduos com alteração/com fenótipo; número de indivíduos com alteração/sem fenótipo.

Foram identificados 5 CI com uma alteração no gene *PCSK9*. Dois CI não relacionados, recrutados durante o período A com a mutação p.Asp374His no exão 7 (Bourbon *et al.* 2008), tendo sido posteriormente comprovada a sua patogenicidade (Fasano *et al.* 2009). Identificaram-se, também, 3 possíveis alterações p.Ala62Asp, p.Arg237Trp e p.Pro467Ala nos exões 1, 5 e 9 respetivamente. A alteração p.Ala62Asp foi encontrada em dois CI, num dos quais esta alteração aparece em heterozigotia composta com a alteração p.Pro467Ala. No decorrer do estudo familiar verificou-se que ambos os pais apresentam valores de colesterol total e LDL de diagnóstico clínico de FH (Figura III.1.1) dado que a mãe apresenta a alteração p.Ala62Asp e o pai a alteração p.Pro467Ala. Não foi possível realizar estudos funcionais para as alterações p.Ala62Asp e p.Pro467Ala, existindo apenas estudos funcionais para a alterações p.Arg238Trp (Benjannet *et al.* 2004; Berge *et al.* 2006; Cameron *et al.* 2006; Kotowski *et al.* 2006; Homer *et al.* 2008). Contudo, a presença de todas as alterações foi verificada num painel de 95 normolipidémicos portugueses não tendo sido encontrada nenhuma das alterações neste painel. Foi também efetuada uma análise *in silico* para todas as

alterações através de 6 programas bioinformáticos, que tiveram por base o efeito da alteração na proteína e na conservação entre espécies do aminoácido e nucleótido alterados, classificando as alterações p.Ala62Asp, p.Arg237Trp, e p.Pro467Ala como possivelmente patogénica, provavelmente patogénicas e neutra, respetivamente (Tabela III.1 1.3). A alteração p.Arg237Trp foi encontrada apenas em um CI. O estudo familiar, foi possível num familiar apenas, verificando-se que o familiar hipercolesterolémico não apresentava esta alteração.

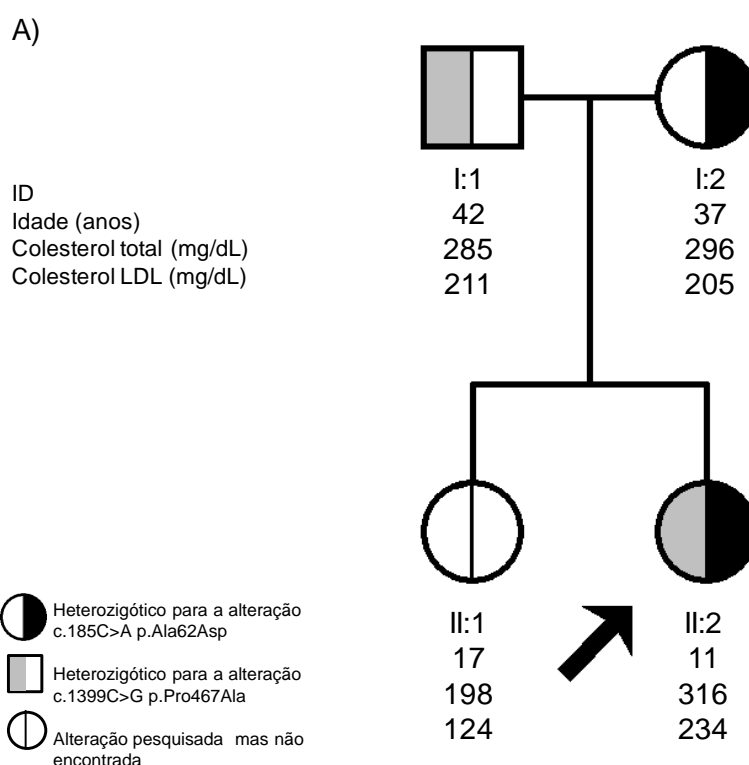


Figura III.1.1 – Caso índice11297 (II:2) com alteração c.185C>A p.Ala62Asp e c.1399C>G p.Pro467Ala no gene PCSK9 (A) Árvore genealógica. A seta representa o caso índice. (B) Sequência genómica de parte do exão 1 do caso índice e seus familiares. A alteração c.185C>A p.Ala62Asp encontrou-se no CI (II:2) e na mãe (I:2) (C) Sequência genómica de parte do exão 9 do caso índice e seus familiares. Observou-se a presença da alteração c.1399C>G p.Pro467Ala no CI (II:2) e no pai (I:1).



Tabela III.1.2 – Alterações encontradas no gene PCSK9 e sua frequência nos CI estudados.

Localização	Alteração		Frequência na população estudada (%)	Fenótipo associado/ dbSNP/ Referência
	cDNA	Proteína		
Promotor	c.1-64 C>T	p.?	29,0	Hipocolesterolemia (Abifadel <i>et al.</i> 2003)
1	c.61_63Ins CTG	p.21_22InsLeu	36,7	Hipocolesterolemia (Yue <i>et al.</i> 2006)
1	c.137 G>T	p.Arg46Leu	1,0	Hipocolesterolemia rs11591147 (Cameron <i>et al.</i> 2006)
1	c.141 C>T	p.Ser47Ser	9,6	Hipocolesterolemia rs28385701 (Abifadel <i>et al.</i> 2009)
1	c.158 C>T	p.Ala53Val	9,2	Hipocolesterolemia rs11583680 (Kotowski <i>et al.</i> 2006)
1	c.207+15G>A	p.?	13,8	rs2495482
Intrão 4	c.657+9 G>A	p.?	12,3	rs11800243
Intrão 4	c.657+76 A>G	p.?	14,2	rs11806638
Intrão 4	c.657+82 A>G	p.?	55,0	rs625619
Intrão 4	c.657+114insA	p.?	46,0	rs112096465
Intrão 4	c.658-7C>T	p.?	65,5	rs2483205
Intrão 5	c.799+3G>A	p.?	81,1	rs2495477
7	c.1026G>A	p.Gln342Gln	14,5	rs509504 (Abifadel <i>et al.</i> 2009)
9	c.1380A>G	p.Val460Val	29,8	rs540796 (Abifadel <i>et al.</i> 2009)
9	c.1420G>A	p.Ile474Val	29,8	Ambos rs562556 (Kotowski <i>et al.</i> 2006)
12	c.2009A>G	p.Glu670Gly	9,5	Ambos rs505151 (Abifadel <i>et al.</i> 2003)

Tabela III.1.3 – Possíveis alterações causadoras de FH encontradas no gene PCSK9.

Exão	Alteração		Co-segreg.	Panel Normolip	Análise in silico				phyloP	Previsão in silico *	Estudo funcional	Referencia	
	cDNA	Proteína			Polyphen2	SIFT	Condel	Mutation Taster					Grantham Score
1	c.185C>A	p.Ala62Asp	2/2‡;0/1	0/95	Possível	Patogénica	Neutra	Causa doença	Mod. Radical	2.38	Possível	Não realizado	Nova
5	c.709C>T	p.Arg237Trp	1/2; 0/0	0/95	Provável	Patogénica	Patogénica	Causa doença	Mod. Radical	0.29	Provável	Não patogénica	(Benjannet et al. 2004; Cameron et al. 2006;)
7	c.1120G>C	p.Asp374His	3/3; 0/3	0/95	Provável	Patogénica	Patogénica	Causa doença	Mod. Conserv.	4.89	Provável	Patogénica	(Bourbon et al. 2008)
9	c.1399C>G	p.Pro467Ala	2/2‡;0/1	0/95	Possível	Neutra	Neutra	Polimorfismo	Conservativa	0.85	Neutra	Não realizado	Nova

Co-segregação: com alteração/com fenótipo; com alteração/sem fenótipo; Nomolip, normolipidémicos; Mod. Conserv., moderadamente conservada; \* Previsão *in silico* com base em 6 programas bioinformáticos considerados, como descrito no Capítulo II. ‡ a alteração encontra-se na mãe (hipercolesterolémica) e no caso-índice (heterozigotia composta p.Ala62Aps + p.Pro467Ala); † a alteração encontra-se no pai (hipercolesterolénico) e no caso-índice (heterozigotia composta p.Ala62Aps + p.Pro467Ala)

### III.1.3 Pirosequenciação

#### III.1.3.1 *454 Life Science*

Em 2009 quando as novas plataformas de *next generation sequencing* (NGS) começaram a ficar disponíveis em Portugal, foram selecionados 65 casos índice (49 adultos e 16 crianças) para o estudo completo dos genes *APOB* e *LDLR* através da técnica de pirosequenciação. Estes CI tinham sido estudados previamente, não tendo sido encontradas mutações no gene *LDLR*, nos 2 fragmentos do gene *APOB* contendo parte dos exões 26 e 29, usualmente analisados no diagnóstico da FH, ou nos 12 exões do gene *PCSK9*. As características clínicas e bioquímicas dos CI estudados por pirosequenciação encontram-se na tabela III.1.4.

Todos os exões dos genes *LDLR* e *APOB* foram amplificados num *pool* de 65 amostras de DNA (20+20+25) e sequenciados pelo método de pirosequenciação (*454 Life Science*, Roche®). As sequências obtidas por pirosequenciação foram alinhadas com base em sequências *consensus* utilizando o *software* Amplicon Variant Analyzer (AVA, Roche). A técnica de pirosequenciação foi realizada pela empresa Biocant em regime de prestação de serviço.

No total obteve-se para o gene *LDLR* um total de 89558 *reads de* nucleótidos correspondendo a uma cobertura média de 30X/fragmentos/indivíduo. Para o gene *APOB* obteve-se um total de 227688 *reads de* nucleótidos correspondendo a uma cobertura média de 35X/fragmentos/indivíduo.

Os resultados obtidos são apresentados na forma de estimativa de alelos. Foram considerados para análise todos os resultados cuja estimativa de alelos apresentavam um valor entre 0,5 e 2,0 (interpretado como presença da alteração em pelo menos um alelo). Uma vez que a técnica foi realizada com base num *pool* de amostras foi necessário confirmar todas as alterações em todas as amostras por sequenciação de Sanger de modo a identificar qual o CI que tinha a alteração.

Após a análise dos resultados referente ao gene *LDLR* verificou-se que o CI 26041 apresentava a mutação c.313+1G>T no intrão 6 e que o CI 29012 apresentava a alteração p.Ala431Thr. As mutações não tinham sido detetadas anteriormente por sequenciação de Sanger na rotina do laboratório do EPHF por motivos técnicos. No caso da mutação no intrão 6 houve uma falha técnica na leitura do eletroforetograma originado pela sequenciação de Sanger. A mutação p.Ala431Thr no exão 9 não foi encontrada no CI 26041 uma vez que os *primers* utilizados para amplificação deste

fragmento continham um SNP que impedia a hibridação normal do alelo com o SNP. Verificou-se que o *primer* reverso hibridava numa zona onde existia o SNP rs6413505. O CI com a mutação p.Ala431Thr era heterozigótico para este SNP, estando ambas as alterações no mesmo alelo dificultando a hibridação do *primer* reverso originando uma menor amplificação deste, não sendo possível detetar a mutação por sequenciação do fragmento amplificado.

As restantes alterações detetadas no gene *LDLR* eram polimorfismo já descritos anteriormente na literatura e que tinham sido previamente identificadas na fase I do EPHF.

Tabela III.1.4 – Caracterização clínica e bioquímica dos 65 CI estudados por pirosequenciação

	Pediátrico n=16	Adultos n=49
Idade (anos)	9,47±3,7	43±12,3
Masculino (%)	50%	63%
CT (mg/dL)	281±84	322±49
cLDL (mg/dL)	201±80	232±45
cHDL (mg/dL)	59±12	55±16
TG (mg/dL)	88±28	140±30
apoA-I (mg/dL)	145±22	144±42
apoB (mg/dL)	127±92	126±44
Lp(a) (mg/dL)	93±100	78±41
Xant Tendon (%)	0%	0%
DCVp (%)	0%	13%
Medicação (%)	12,5%	35%

CT, colesterol total; cLDL, colesterol LDL; cHDL, colesterol HDL; TG, triglicéridos; apoA-I, apolipoproteína A-I; apoB, apolipoproteína B; Lp(a), lipoproteína (a); Xant Tendon, xantomas tendinosos; DCVp, doença cardiovascular prematura.

Na análise aos dados obtidos para o gene *APOB* foram detetadas 87 alterações. Mais de metade eram SNPs descritos anteriormente, dado que 32 eram possíveis alterações patogénicas. Um total de 15 fragmentos diferentes localizados nos exões 13, 14, 15, 16,

19, 22, 23, 24, 26 e 29 (embora a maioria nos exões 26 e 29) foram resequenciados por sequenciação de Sanger.

Das 32 possíveis alterações identificadas por pirosequenciação, apenas 7 foram detetadas por sequenciação de Sanger. Durante a resequenciação foram também detetadas 3 alterações (2 não descrita e uma já descrita) que não tinham sido reportadas na pirosequenciação. As 10 alterações encontram-se descritas na tabela III.I 1.5 encontrando-se sublinhado as alterações que não fora detetadas pela técnica de pirosequenciação. Duas alterações não co-segregavam com a hipercolesterolemia na família e para 2 outras, não foi possível efetuar o estudo de co-segregação familiar (uma vez que os familiares não se mostraram disponíveis para participar). As restantes 6 alterações encontradas co-segregavam com o fenótipo de FH nas famílias, apesar de existirem alguns casos, em especial em familiares mais jovens, com baixa penetrância (Figura III.1.2). Os estudos funcionais de três destas alterações serão apresentados na parte 2 do capítulo III.

Um painel de 95 controlos normolipidémicos portugueses foi analisado quanto à presença das possíveis 10 alterações. As alterações p.Asp1113His, p.Tyr1247Cys e p.Ser3274Gly foram encontradas em 2 indivíduos normolipidémicos e a alteração p.Asp2213del num indivíduo do painel normolipidémico, por esta razão foram consideradas polimorfismos (Figura III.1.2).

As 10 alterações no gene *APOB* foram também pesquisadas num painel de 95 indivíduos negativos para a FH e apenas as alterações p.Asp1113His e p.Ser3274Gly estavam presentes em 3 e 2 indivíduos, respetivamente. Estas alterações estão descritas na base de dados de SNP (<http://www.ncbi.nlm.nih.gov/SNP/>) com uma frequência de 0,3% e 0,7%, respetivamente (Figura III.1.2).

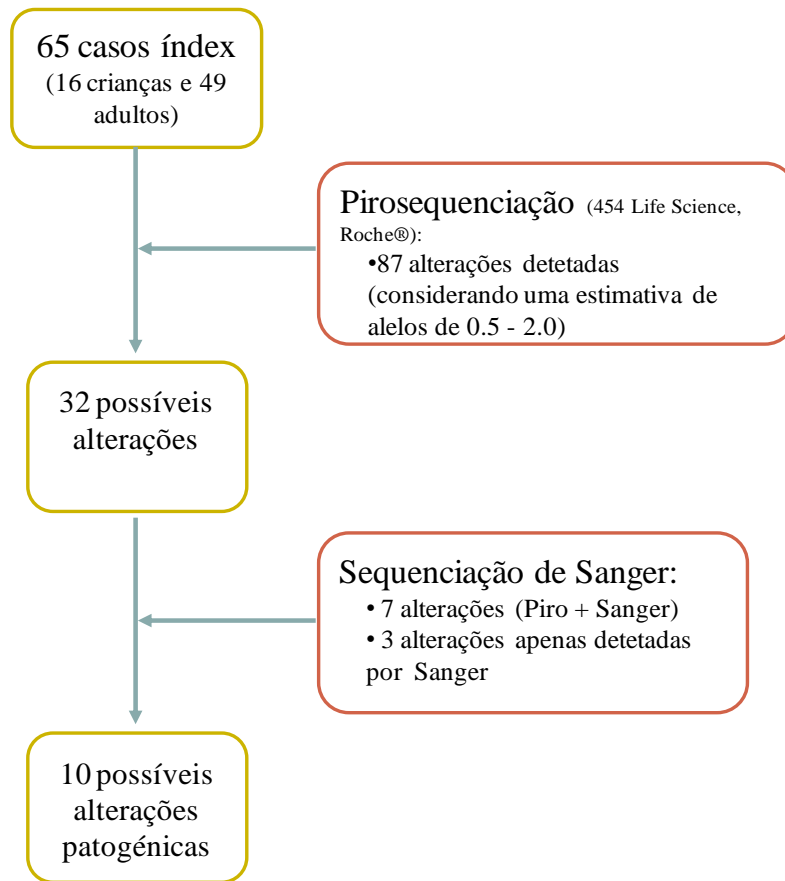


Figura III.1.2 – Diagrama esquemático do estudo por pirosequenciação e resequenciação de Sanger de 65 CI do EPHF.

Tabela III.1.5 – Características clínicas dos CI com alterações detetadas por pirosequênciação no gene *APOB*

CI	Idade	Alteração		CT (mg/dL)	cLDL (mg/dL)	cHDL (mg/dL)	TG (mg/dL)	apoB (mg/dL)	apoA-I (mg/dL)	Lp(a) (mg/dL)	DCV (idade)	dbSNP
		Exão	cDNA	Proteína								
29012	59	19	c.2981C>T	p.Pro994Leu	302	230	49	113	87*	130*	74*	EAM (56) rs41288783
26078	14	22	c.3337G>C	p.Asp1113His	217	140	73	34	65,8	117	83,6	No rs12713844
25023	58	22	c.3491G>C	p.Arg1164Thr	350	255	56	195	81,3*	178*	132*	No nd
28030	41	24	c.3740A>G	p.Tyr1247Cys	251*	141*	36*	372*	136,5*	120*	240*	EAM (40) rs61741164
28051	44	26	c.5599C>T	p.Arg1867Trp	349	255	46	261	185	167	85	No rs200583769
27058	7	26	c.5690G>A	p.Arg1897His	241	189	43	129	69	124	170	No rs199510126
28171	66	26	c.6639_6641delTGA	p.Asp2213del	352	246	55	255	93*	175*	130*	EAM (50) rs72653087
28124	34	26	c.7853T>C	p.Ile2618Thr	>350	>200	51*	100	93*	180*	<8*	No nd
28086	11	26	c.9835A>G	p.Ser3279Gly	280	207	57	78	67*	152*	143*	No rs12720854
28120	21	29	c.13480_13482delCAG	p.Gln4494del	256	192	33	157	119	109	<8	No nd

CI, caso-índice; dbSNP – base de dados SNP; \* em medicação; EAM, enfarte agudo do miocárdio (anos); nd – não descrito; a sublinhado encontram-se as alterações detetadas apenas por sequenciação de Sanger

### III.1.2.2 *GS Junior System*

Em 2012 foram selecionados 45 CI com diagnóstico clínico de FH para o estudo por pirosequenciação no equipamento *GS Junior system*. Foi desenvolvido um método de diagnóstico para a FH, que cobria todos os exões e uma parte dos intrões para os genes *LDLR*, *APOB*, *PCSK9* e *LDLRAP1*, simultaneamente, pela empresa Progenika. A seleção dos 45 CI teve em conta os indivíduos que apresentavam um fenótipo agressivo de FH e não tinha nenhuma alteração patogénica num dos três genes estudados no âmbito do EPHF.

Uma vez que os resultados anteriores de pirosequenciação através do equipamento *454 Life Science* originou um grande número de falsos positivos, bem como, falsos negativos, juntamente com as 45 amostras de CI sem alteração num dos três genes estudados no âmbito do EPHF, foram enviadas 12 amostras controlos com uma mutação identificada nos genes *LDLR*, *APOB* e *PCSK9*. O estudo foi realizado a 35 CI adultos e 10 CI crianças (Tabela III.1.7). A técnica de pirosequenciação foi realizada pela empresa Progenika, em regime de prestação de serviço.

Após o estudo molecular verificou-se que em todos os controlos enviados foram identificadas por este método todas as alterações nos genes *LDLR*, *APOB* e *PCSK9*, verificando-se uma melhoria significativa da técnica. Apenas o gene *LDLRAP1* não tinha sido estudado no âmbito do EPHF, em todas as amostras enviadas, CI e controlos. Das 57 amostras estudadas foram detetadas as alterações p.Ile132Val, p.Thr218Ile, p.Arg238Trp e p.Asp293His em heterozigotia neste gene. Após estudo familiar verificou-se que para as alterações p.Ile132Val, p.Thr218Ile e p.Arg238Trp apenas os CI apresentavam as alterações, uma vez que os familiares com hipercolesterolemia não apresentavam estas alterações, desta forma não se verificou a co-segregação da alteração com o fenótipo de hipercolesterolemia. No caso da alteração p.Asp293His, verificou-se que estava presente no CI, pai, tio-avó paterno e primos do lado paterno, todos com hipercolesterolemia (Figura III.1.3). A alteração p.Arg238Trp encontra-se descrita como rs4129105, apresentado uma frequência de 2,4% na base de dados de SNP (<http://www.ncbi.nlm.nih.gov/SNP/>). A alteração p.Thr218Ile, encontra-se também descrita na mesma base de dados com a referência rs114583297 e uma frequência de 0,6%, não existindo estudos funcionais a comprovarem, ou não, a sua patogenicidade. As alterações p.Ile132Val e p.Asp293His não se encontram descritas





Tabela III.1.7 – Caracterização clínica e bioquímica dos 43 CI estudados pelo equipamento *GS Junior System*.

	<b>Pediátrico</b>	<b>Adultos</b>
	<b>n=10</b>	<b>n=35</b>
<b>Idade (anos)</b>	11±4	48±11
<b>Masculino (%)</b>	33,30%	60,80%
<b>CT (mg/dL)</b>	285±54	360±85
<b>cLDL (mg/dL)</b>	201±49	268±56
<b>cHDL (mg/dL)</b>	60±21	57±13
<b>TG (mg/dL)</b>	75±22	139±49
<b>apoA-I (mg/dL)</b>	142±34	168±36
<b>apoB (mg/dL)</b>	105±26	142±35
<b>Lp(a) (mg/dL)</b>	54±74	58±62
<b>Xantomas (%)</b>	0%	4,30%
<b>pDCV (%)</b>	0%	36,90%
<b>Medicação (%)</b>	11,10%	74%

CT, colesterol total; cLDL, colesterol LDL; cHDL, colesterol HDL; TG, triglicéridos; apoA-I, apolipoproteína A-I; apoB, apolipoproteína B; Lp(a), lipoproteína (a); pDCV, doença cardiovascular prematura.





## **Parte 2 – Caracterização funcional de alterações encontradas nos genes associados a FH no âmbito do Estudo Português de Hipercolesterolemia Familiar**

Este trabalho está incluído nos seguintes artigos:

Bourbon M, **Alves AC**, *et al.* (2008) *Atherosclerosis* 196: 633–642

Bourbon M, Duarte MA, **Alves AC**, *et al.* (2009) *J. Med. Genet.* 46;352-357

Medeiros AM, **Alves AC**, *et al.* (2010) *Atherosclerosis* 212:553–558

Silva S, **Alves AC**, *et al.* (2012) *Atherosclerosis* 225; 128-134

**Alves AC**, *et al.* (2013) *Hum. Mol. Genet* doi:10.1093/hmg/ddt573

**Alves AC**, *et al.* (2014) submitted



## **Parte 2 – Caracterização funcional de alterações encontradas nos genes associados a FH no âmbito do Estudo Português Hipercolesterolemia Familiar**

Entre os anos de 1999 e 2012 foram identificados do ponto de vista molecular 294 CI no EPHF, correspondendo a 117 alterações diferentes no gene *LDLR*, 39 das quais foram descritas pela primeira vez na população portuguesa (Bourbon *et al.* 2008; Medeiros *et al.* 2010, Alves *et al.* 2014 submitted). As alterações identificadas no gene *LDLR* em doentes com FH incluem 57 mutações *missense*, 25 mutações *nonsense*, 6 grandes rearranjos, 2 alterações na região promotora do gene, 19 mutações de *splicing* e 8 substituições sinónimas. No gene *APOB* foram encontradas 13 alterações diferentes e no gene *PCSK9* 4 alterações diferentes.

Apesar de terem sido identificadas 134 alterações nos 3 genes estudados, apenas foram realizados estudos funcionais para comprovar a patogenicidade de 32 destas alterações, o que corresponde a que somente 175 CI dos 294 apresentem uma mutação comprovadamente patogénica e causadora de doença. De forma a comprovar a patogenicidade das alterações de *splicing*, promotor e sinónimas, foram realizados estudos funcionais por RT-PCR a partir de mRNA extraído dos casos índice com as alterações em estudo. Para as alterações *missense* os estudos funcionais foram realizados a partir de linfócitos dos casos índice com as alterações no gene *LDLR*; para as alterações no gene *APOB* os estudos foram realizados a partir das partículas de LDL, isoladas do soro dos indivíduos com as alterações em estudo.

### **III.2.1 Estudo funcional de alterações no promotor**

Uma vez que mutações no promotor afetam a expressão de todo o gene foram realizados estudos funcionais de forma a comprovar a patogenicidade da alteração c.-13A>G, descrita apenas na população portuguesa e para a qual não existem até à data estudos funcionais. De forma a comprovar a patogenicidade da alteração foi realizado o estudo por RT-PCR a partir do mRNA do CI, com o objetivo de verificar a presença do polimorfismo no exão 2 (rs2228671) que tinha sido detetado no DNA genómico do CI ou se a alteração na região do promotor originava a não expressão do alelo com a alteração. Após o estudo verificou-se que a alteração não é patogénica uma vez que foi

possível visualizar o polimorfismo rs2228671 no cDNA, tal como se observava no DNA genómico, ou seja, ambos os alelos estavam a ser expressos (Figura III.2.1). O resultado *in silico* das alterações encontradas na região do promotor encontram-se na tabela III.2.2 juntamente com as alterações de *splicing*.

A)

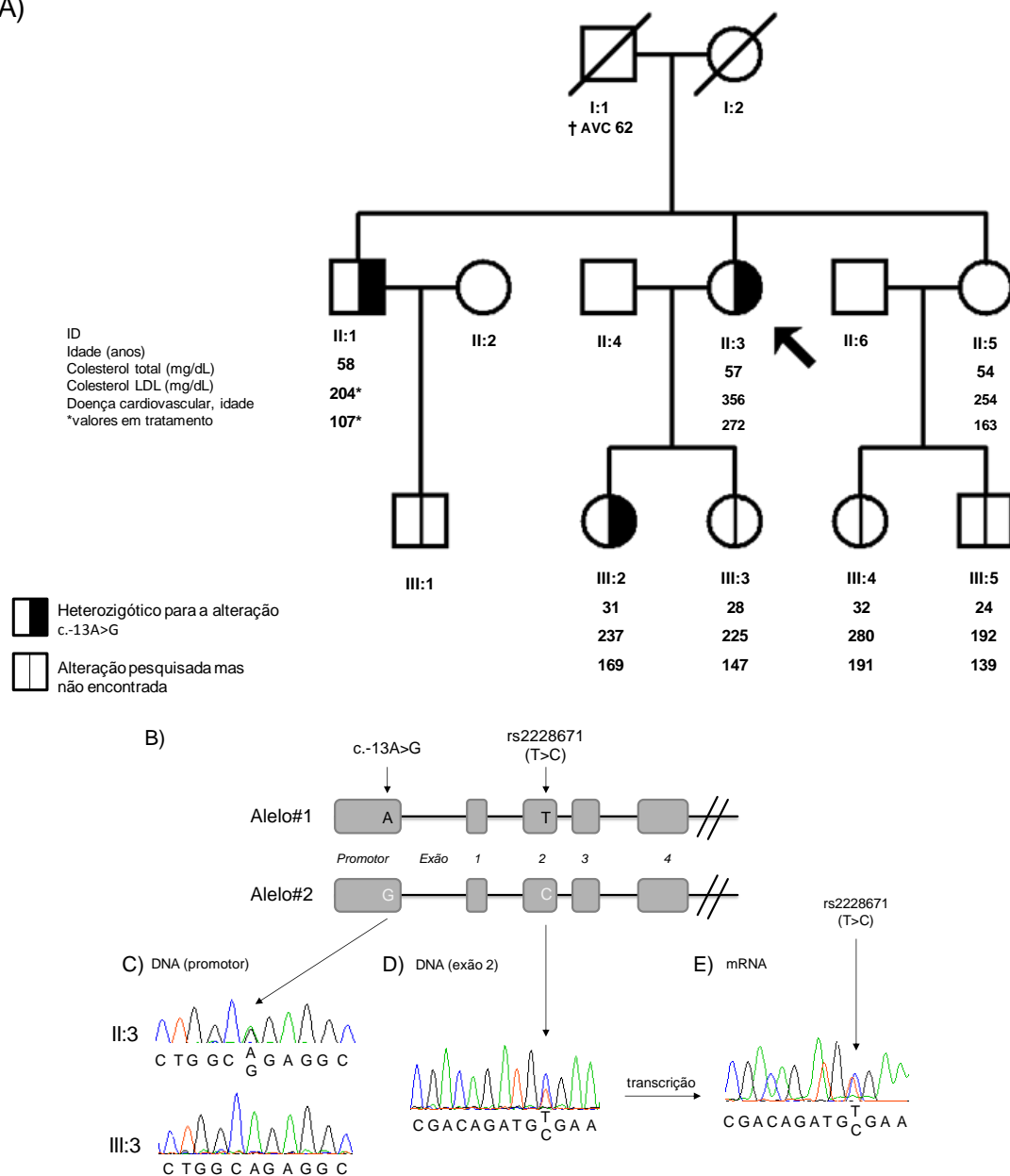


Figura III.2 1 – Estudo funcional do caso índice 29105 (II:3) com a alteração c.-13A>G no gene *LDLR*. (A) Árvore genealógica. A seta representa o caso índice. AVC, acidente vascular cerebral; † morte. (B) Representação esquemática da localização da alteração c.-13A>G no promotor e do polimorfismo rs2228671 (T>C) no exão 2 do gene *LDLR*. (C) sequência genômica de um fragmento contendo parte do promotor do *LDLR* do caso índice e de um familiar, onde se observa a alteração c.-13A>G apenas no CI. (D) Sequência nucleótica de um fragmento contendo parte do exão 2, a partir de DNA genômico do CI, onde se observa o polimorfismo rs2228671. (E) Sequência nucleótica de um fragmento contendo parte do exão 2, (mRNA do CI) onde se observa também o polimorfismo rs2228671.



### III.2.2 Estudo funcional de alterações de *splicing*

Foram encontradas 19 alterações de *splicing* no gene *LDLR* em 38 CI, sendo 12 exclusivas da população portuguesa. Das 19 alterações de *splicing* encontradas durante este trabalho, 5 tinham sido já analisadas do ponto de vista funcional (Hobbs *et al.* 1992; Leren *et al.* 1994; Lombardi *et al.* 1995; Mak *et al.* 1998). De forma a comprovar a patogenicidade de 11 alterações de *splicing* encontradas nos CI do EPHF, foram realizados estudos de RT-PCR a partir do mRNA dos CI com as alterações c.190+4insTG, c.313+6T>C, c.818-3C>G, c.818-2A>G, c.1060+1G>A, c.1061-8T>C, c.1359-5C>G, c.1706-10G>A, c.2140+5G>A, c.2389G>T e c.2547+1G>A. Verificou-se que as alterações c.190+4insTG, c.313+6T>C, c.818-3C>G, c.818-2A>G, c.1060+1G>A, c.1359-5C>G, c.2389G>T e c.2547+1G>A são patogénicas produzindo diversas alterações no mRNA (como descrito nas Figuras III.2.2, 2.3 e na Tabela III.2.1) e que as alterações c.1061-8T>C, c.1706-10G>A e c.2140+5G>A (Figura III.2.4) não afetam o correto processamento do pré-mRNA (Tabela III.2.1), não sendo por esta razão causadoras de doença. As alterações em que foram realizados os estudos funcionais durante este trabalho encontram-se apresentadas na tabela III.2.1.

Tabela III.2.1 – Efeito das mutações *splicing* estudadas por RT-PCR no decorrer deste trabalho.

Alteração		Efeito no RNA	Posição do novo codão <i>stop</i> originado	Estudo funcional
Intrão	cDNA			
2	c.190+4insTG	Retenção de 2 nucleótidos do intrão 2	206	Patogénica
3	c.313+6T>C	<i>Skipping</i> exão 3	820	Patogénica
5	c.818-3C>G	Retenção de 2 nucleótidos do intrão 5	274	Patogénica
5	c.818-2A>G	Retenção de 10 nucleótidos do intrão 5	303	Patogénica
7	c.1060+1G>A	<i>Skipping</i> exão 7	779	Patogénica
7	c.1061-8T>C	Nenhum	---	Não patogénica
9	c.1359-5C>G	Retenção do intrão 9	454	Patogénica
11	c.1706-10G>A	Nenhum	---	Não patogénica
14	c.2140+5G>A	Nenhum	---	Não patogénica
16	c.2389G>T (p.Val797Leu)	<i>Skipping</i> exão 16	835	Patogénica
17	c.2547+1G>A	<i>Skipping</i> exão 17	805	Patogénica

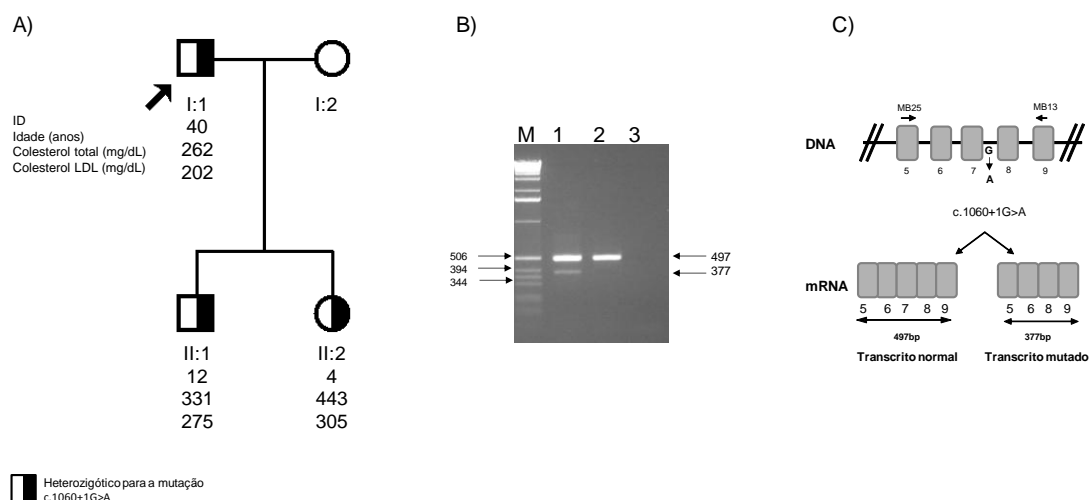


Figura III.2.2 – Estudo funcional do caso índice 26110 (I:1) com a mutação c.1060+1G>A. (A) Árvore genealógica. A seta representa o caso índice. (B) Eletroforese em gel de agarose dos produtos de RT-PCR de mRNA do CI e de um indivíduo normal, amplificados com os *primers* localizados no exão 5 (MB25) e exão 9 (MB13). M, marcador; 1 caso índice; 2 controlo; 3 controlo sem cDNA. (C) representação esquemática da localização dos *primers* utilizados e dos 2 fragmentos obtidos.

Para as alterações c.1586+2T>A e c.1845+1delG não foi possível obter amostra e realizar os estudos funcionais, tendo-se optado assim por uma análise *in silico* com base no seguintes programas bioinformáticos, *Splice-Site Predictor (Splice Port)*, *Neural Network Splice Site Prediction Tool (NNSSP)* e *neural network predictions of splice sites in human (NetGen2)*, que será apresentada posteriormente. Os 12 CI onde a patogénica da alteração encontrada não foi comprovada foram considerados como não tendo uma alteração causadora de doença num dos três genes associados à FH.

### III.2.3 Estudo funcional de alterações sinónimas

Em 2007 Bourbon *et al* identificaram a primeira alteração silenciosa patogénica no gene *LDLR*, facto que despertou o interesse no estudo de todas as mutações silenciosas no gene *LDLR*. Durante este trabalho foram encontradas 8 alterações sinónimas em 18 CI.

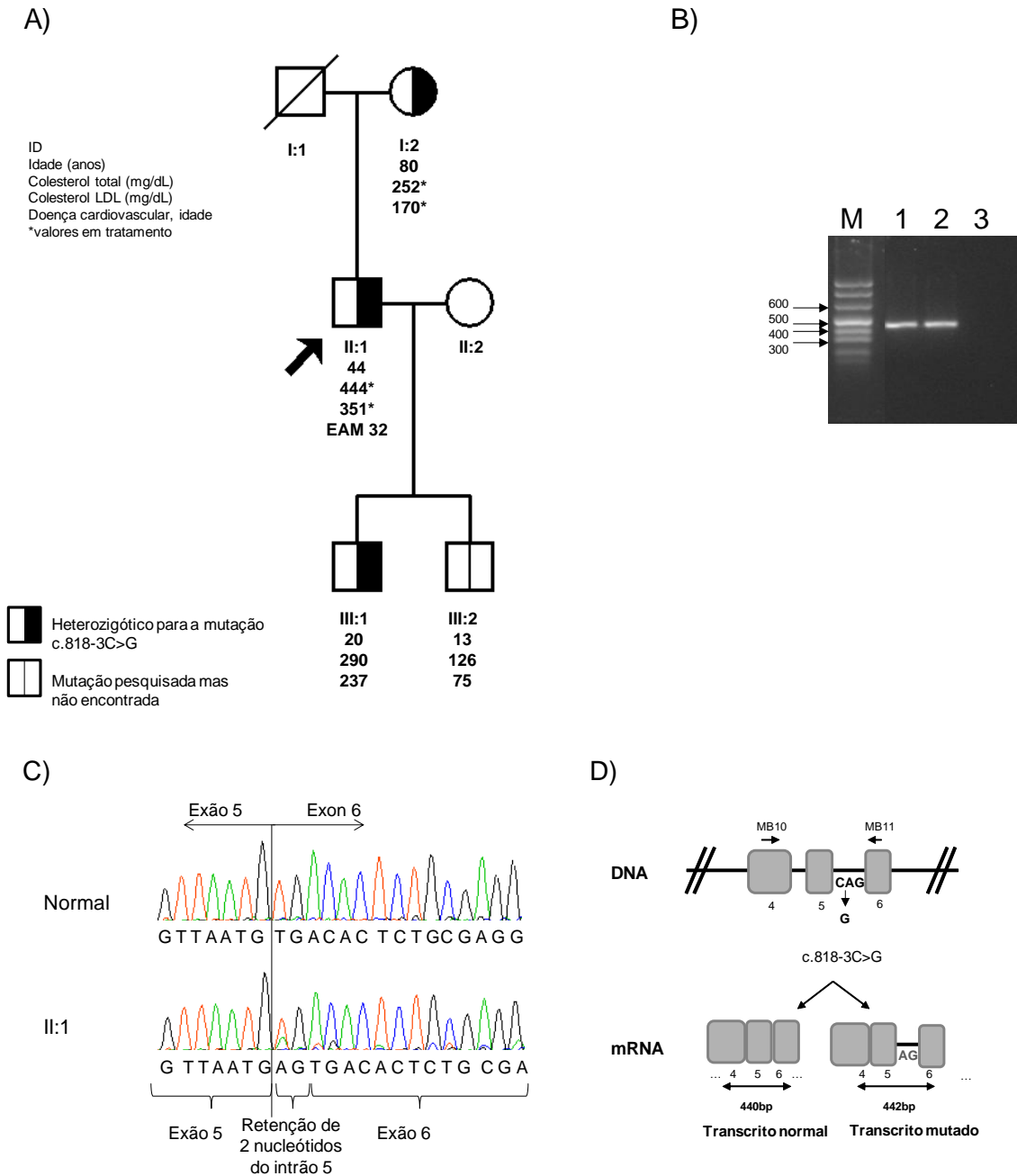


Figura III.2.3 – Estudo funcional do caso índice 12003 (II:1) com a mutação c.818-3C>G. (A) Árvore genealógica. A seta representa o caso índice. EAM, enfarte agudo do miocárdio. (B) Eletroforese em gel de agarose dos produtos de RT-PCR de mRNA do CI e de um indivíduo normal, amplificados com os *primers* localizados no exão 4 (MB10) e exão 6 (MB11). M, marcador; 1 caso índice; 2 controle; 3 controle sem cDNA. (C) Sequência nucleótida de parte do fragmento amplificado do cDNA do CI onde se verificou a retenção de 2 nucleótidos do intrão 5 (D) representação esquemática da localização dos *primers* utilizados e dos 2 fragmentos obtidos

Foram desenvolvidos estudos funcionais por RT-PCR a partir do mRNA dos CI com as seguintes alterações: c.1911C>T, c.1920C>T e c.1977C>T. Após o estudo funcional verificou-se que as alterações não afetam o correto processamento do pré-mRNA (Figura III.2.5). A alteração sinónima c.1185G>C foi encontrada num CI que apresentava também a alteração c.-135G>C, neste caso, o estudo RT-PCR revelou que a alteração no promotor originava a não expressão do alelo e que a alteração sinónima c.1185G>C se encontrava no mesmo alelo (Bourbon *et al.* 2009). Para as restantes alterações não foi possível realizar o estudo funcional uma vez que não foi possível obter nova amostra dos CI. Os 3 CI cujas alterações não são patogénicas foram classificados como não tendo FH.

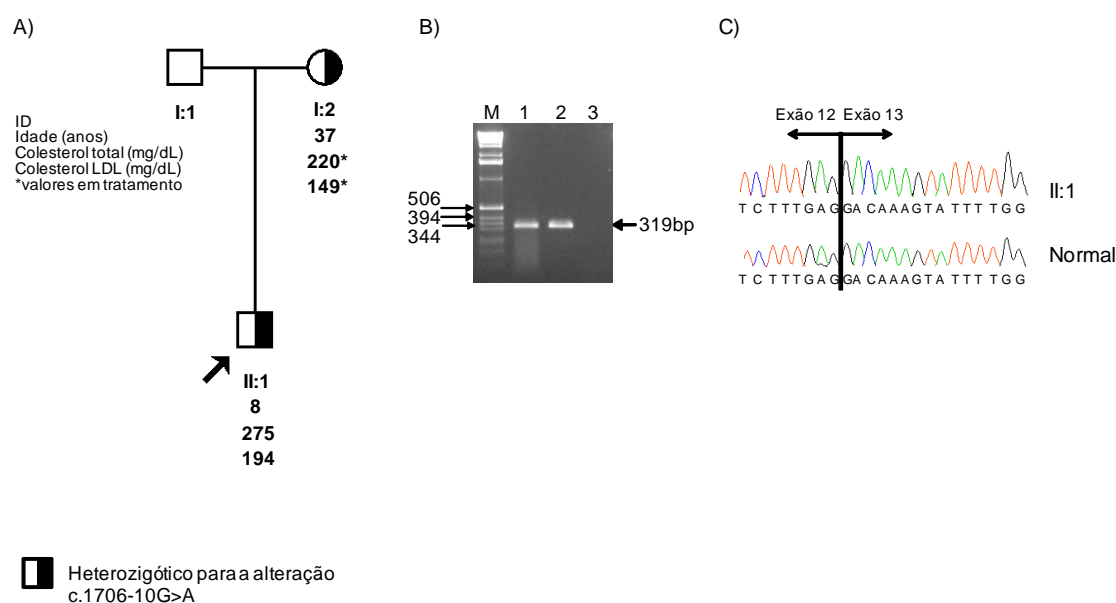


Figura III.2.4 – Estudo funcional do caso índice 26082 (II:1) com a alteração c.1706-10G>A. (A) Árvore genealógica. (B) Eletroforese em gel de agarose dos produtos de RT-PCR de mRNA do CI e de um indivíduo normal, amplificados com os *primers* localizados no exão 11 (MB16) e exão 13 (MB17). M, marcador; 1 caso índice; 2 controlo normal; 3 controlo sem cDNA. (D) Sequência de nucleótidos de parte do fragmento amplificado do controlo normal e do CI, apresentado, ambas as amostras, o exão 12 e 13 do *LDLR*, não se verificando qualquer efeito desta alteração no processamento do RNA.

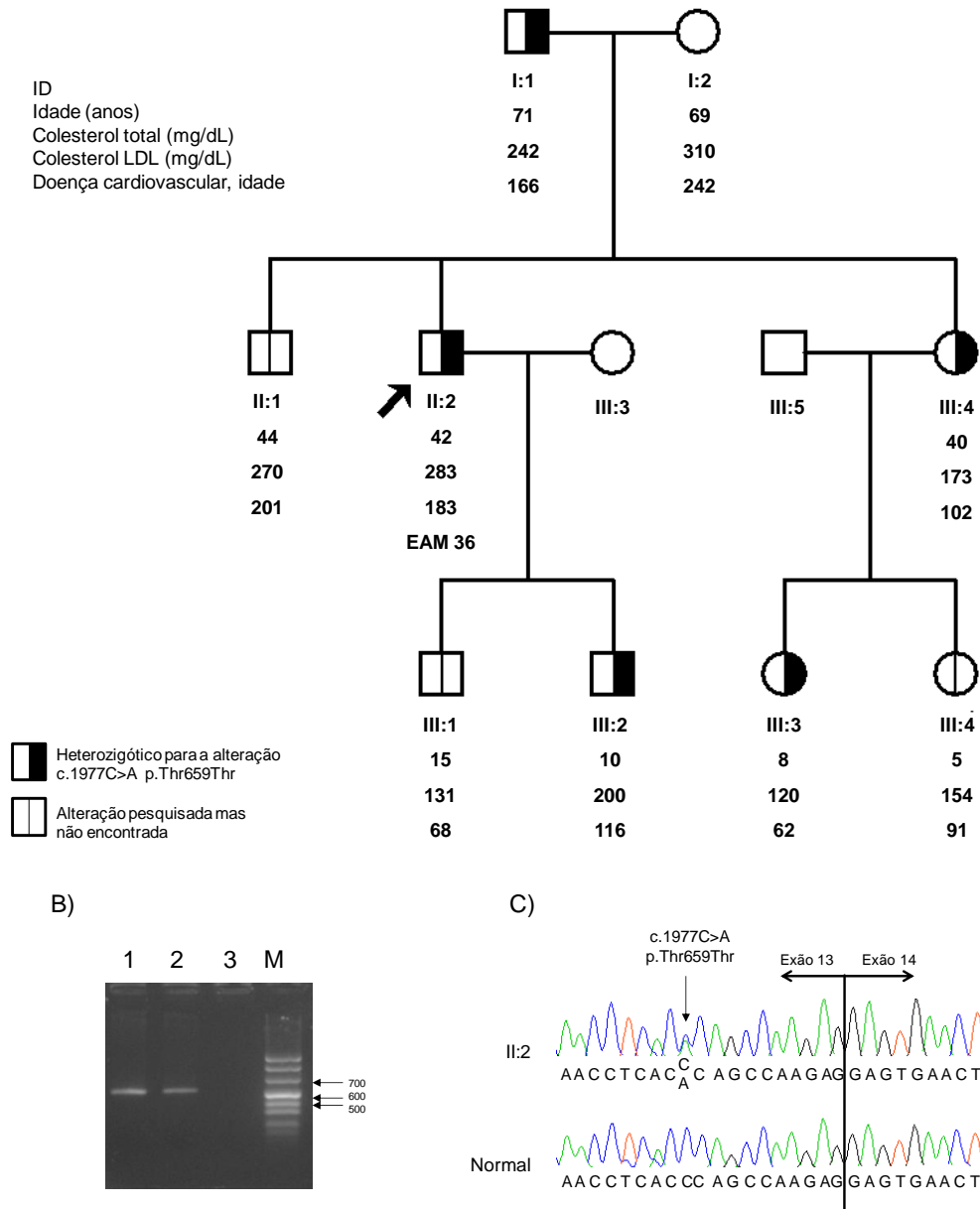


Figura III.2.5 – Estudo funcional do caso índice 25024 (II:2) com a alteração p.Thr659Thr. (A) Árvore genealógica. (B) Eletroforese em gel de agarose dos produtos de RT-PCR de mRNA do CI e de um indivíduo normal, amplificados com os *primers* localizados no exão 11 (MB16) e exão 15 (MB19). M, marcador; 1 caso índice; 2 controle normal; 3 amostra sem cDNA. (D) Sequência de nucleótidos de parte do fragmento amplificado do controle normal e do CI, apresentado, ambas, o exão 13 e 14 do *LDLR* não se verificando qualquer efeito no processamento correto do mRNA.

### III.2.4 Estudo funcional de alterações pontuais no gene *LDLR*

No gene *LDLR* foram encontradas 59 alterações pontuais em 166 CI do EPHF, destas apenas 22 alterações tinham estudos funcionais comprovando a sua patogenicidade (Bertolini *et al.*, 1999; Fouchier *et al.*, 2001; Hobbs *et al.*, 1990, 1992; Leitersdorf *et al.*, 1993; Pereira *et al.*, 1995; Soutar *et al.*, 1989). Foi verificada a presença de todas as alterações não caracterizadas num painel de 95 indivíduos portugueses normolipidémicos, não tendo sido encontrada nenhuma destas alterações.

Para as alterações p.Gly76Trp, p.Ile441Thr, p.Gly545Trp e p.Cys698Phe foram realizados estudos funcionais por citometria de fluxo a partir de linfócitos dos CI com as alterações, como descrito no ponto II.7.2 do capítulo II. Os linfócitos dos CI foram colocados em crescimento em cultura celular e os ensaios foram efetuadas a 4°C para verificar a ligação do recetor das LDL dos CI às partículas de LDL (de indivíduos normais) e a 37°C para os ensaios de internalização. Foram também realizados estudos de expressão a 4°C com anticorpos específicos de modo a verificar a expressão do gene *LDLR*. As partículas de LDL utilizadas nos ensaios foram isoladas a partir de indivíduos normolipidémicos. De forma a validar os ensaios foram realizados estudos com amostras de CI com mutações patogénicas comprovadas funcionalmente: p.Ala431Thr, p.Gly478Arg e Ex3\_Ex4del; bem como amostras de um indivíduos normolipidémicos, considerados como *LDLR wild-type* (wt). Todos os ensaios foram efetuado em triplicado e os resultados apresentados são a média de pelo menos dois ensaios independentes.

O estudo molecular ao CI 25018 identificou a alteração p.Gly76Trp, tendo-se verificado por estudos familiares que esta alteração não co-segrega na família com o fenótipo de FH (Figura III.2.6). Os estudos funcionais realizados verificaram que as células com a alteração p.Gly76Trp apresentavam uma ligação e internalização do complexo LDL:LDLR semelhante ao controlo normolipidémico (Tabela III.2.2 e Figura III.2.7A,B e C) tendo sido esta alteração classificada como neutra.

A alteração p.Ile441Thr foi encontrada em 4 CI, o estudo molecular às famílias dos 4 CI indicaram existir co-segregação (dado que, apenas um familiar hipercolesterolémico de um CI não apresentava a alteração). Após os estudos funcionais verificou-se que a taxa de ligação do complexo LDL:LDLR era de aproximadamente 95% e a internalização de 65%, sendo os valores de expressão do LDLR normais (Tabela III.2.4 e Figura III.2.8A,B e C). Esta alteração foi classificada como patogénica.

A alteração p.Gly545Trp foi encontrada em 8 CI, após o estudo familiar verificou-se a existência de co-segregação nas famílias dos respectivos CI. Os ensaios funcionais mostraram existir uma redução de cerca de 40% na internalização e de aproximadamente 20% na ligação do complexo LDL:LDLR. Os estudos de expressão mostraram que a mesma era normal (Tabela III.2.2 e Figura III.2.9A,B e C). Esta alteração foi classificada como patogénica.

Tabela III.2.2 – Análise da atividade e expressão do LDLR (expressão, ligação e internalização) de 4 alterações pontuais encontradas no *LDLR* e 2 controlos com mutações funcionais.

Alteração	Proteína	Expressão e atividade do LDLR			Estudo funcional	Observações
		Expressão	Ligação	Internalização		
cDNA						
c.1291G>A	p.Ala431Thr	90%	48%	70%	Patogénica	Controlo
c.1432G>A	p.Gly478Arg	95%	55%	48%	Patogénica	Controlo
c.226G>T	p.Gly76Trp	96%	94%	97%	Neutra	Nova
c.1322T>C	p.Ile441Thr	97%	95%	65%	Patogénica	Nova
c.1633G>T	p.Gly545Trp	92%	80%	60%	Patogénica	Nova
c.2093G>T	p.Cys698Phe	93%	50%	50%	Patogénica	Nova

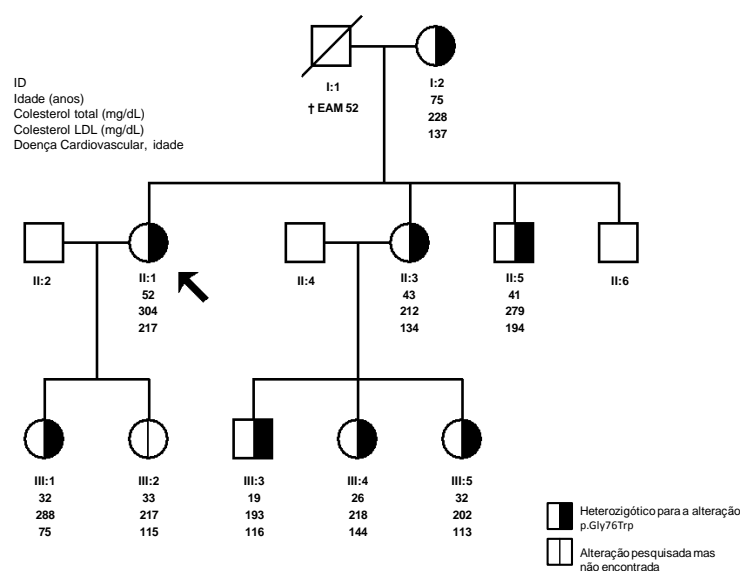
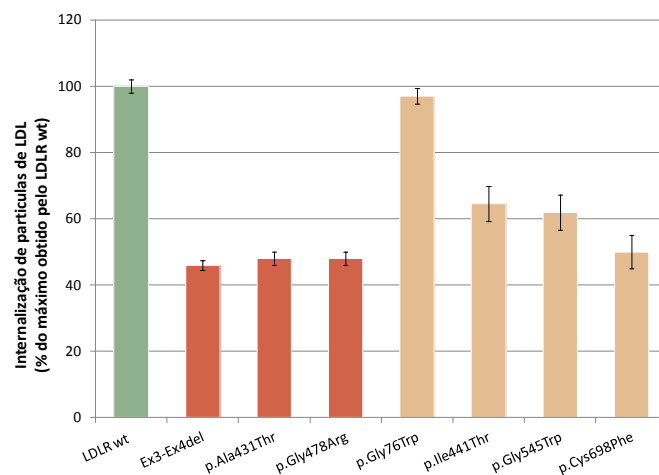
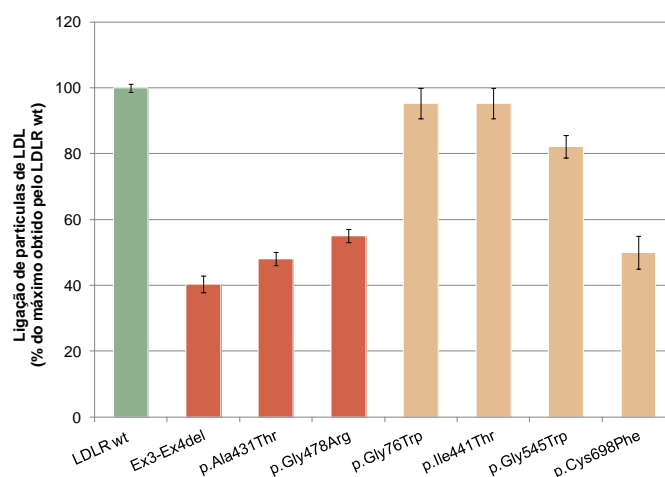


Figura III.2.6 – Árvore genealógica do CI 25018 com a alteração p.Gly76Trp no exão 3 do gene *LDLR*. A seta representa o caso índice. EAM, enfarte agudo do miocárdio; † morte.

A)



B)



C)

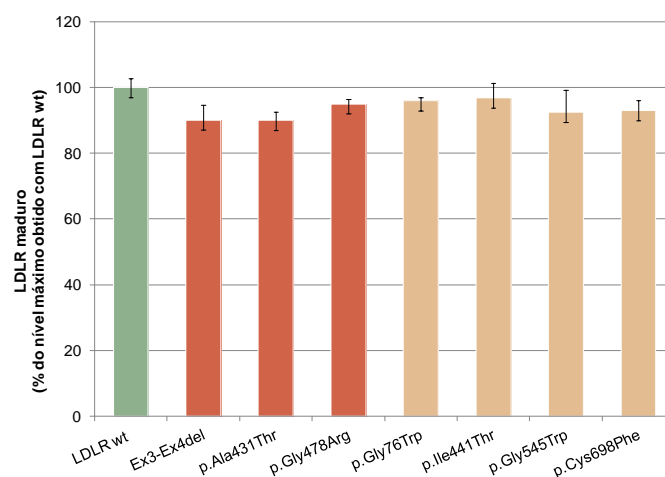


Figura III.2.7 – Níveis de ligação, internalização e expressão dos linfócitos com as diferentes alterações. A) Taxa de ligação das partículas de LDL marcada nos linfócitos após uma incubação de 4 horas a 4°C. B) Taxa de internalização das partículas de LDL marcada nos linfócitos após uma incubação de 4 horas a 37°C LDL. C) Expressão dos linfócitos com LDLR normal (LDLR *wild type*), linfócitos da amostra controlo (Ex3\_Ex4del) e dos linfócitos com as alterações em estudo. \*  $p < 0,05$  entre as alterações de patogenicidade desconhecida e LDLR *wild type* (teste t-sudent).



Os estudos moleculares identificaram 2 CI com a alteração p.Cys698Phe. Também para esta alteração foi observada a sua co-segregação na família (apenas um familiar hipercolesterolémico de um CI não apresentava a alteração). No estudo funcional verificou-se existir uma diminuição de aproximadamente 50% na internalização bem como na ligação do complexo LDL:LDLR. Tal como as alterações anteriores, também os estudos de expressão mostraram uma expressão normal de LDLR das amostras com esta alteração. Esta alteração foi classificada como patogénica.

### III.2.5 Estudo funcional de ligação e internalização em alterações do gene APOB

De forma a verificar a patogenicidade das alterações detetadas, através da técnica de pirosequenciação, no gene *APOB*, isolou-se as partículas de LDL de doentes com as alterações p.Asp1113His, p.Arg1164Thr, p.Tyr1247Cys e p.Gln4494del, bem como de familiares com e sem as alterações, indivíduos normolipidémicos (considerados *wild type* - wt) e indivíduos com a alteração APOB3527 (alteração mais comum no gene *APOB* comprovadamente patogénica) (Figura III.2.8). Os ensaios funcionais foram efetuadas por citometria de fluxo em linfócitos e linhas celulares de HepG2, a partir das partículas de LDL dos CI com as alterações. Foi também verificada a proliferação das células U937 quando encubadas com partículas de LDL dos CI.

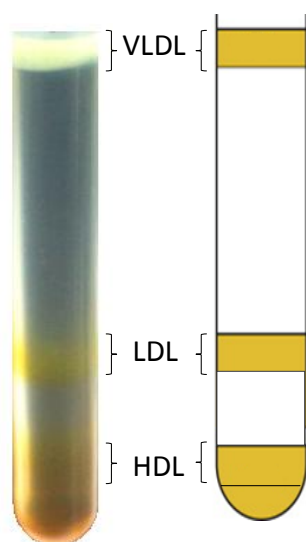


Figura III.2.8 – Preparação das partículas de LDL por ultracentrifugação. Do lado esquerdo encontra-se uma amostra cujo isolamento foi efetuado no laboratório e no lado direito a representação esquemática das partículas obtidas por ultracentrifugação consoante a sua densidade.

A alteração p.Arg1164Thr foi encontrada num CI (25023) e após o estudo familiar verificou-se que co-segregava na família com o fenótipo de FH exceto em dois familiares jovens (Figura III.2.9A). A previsão *in silico* revelou que esta alteração seria possivelmente patogénica, no entanto a previsão *in silico* em proteínas estruturais como a APOB podem não ser fiáveis, sendo extremamente importante o estudo funcional. A alteração não está presente num painel de 95 indivíduos portugueses normolipidémicos, nem na base de SNP (<http://www.ncbi.nlm.nih.gov/snp>) ou *National Heart, Lung, and Blood Institute* (NHLBI) *database* (<http://evs.gs.washington.edu/EVS/>). A alteração foi pesquisada mas não detetada num painel de 95 indivíduos com diagnóstico clínico de FH mas sem alteração identificada até à data (Tabela III.2.5). Os estudos funcionais efetuados em linfócitos *wild type* com as partículas de LDL do CI e 2 familiares com hipercolesterolemia e com a alteração demonstraram existir uma diminuição de cerca de 40% na ligação e internalização quando comparados com o controlo normal (Figura III.2.9B e C). O mesmo resultado foi observado nas células HepG2 e U937 (Figura III.2.9 D e E). De forma a identificar-se a origem do fenótipo mais suave apresentado por alguns dos familiares deste CI fez-se a sequenciação do gene PCSK9 gene onde foram descritas alterações causadoras de hipocolesterolemia bem como a pesquisa da mutação p.Arg19X no gene APOC3 também descrita como causadora de hipocolesterolemia. Nenhuma das alterações foi detetada.

No CI 28120 foi encontrada a alteração p.Gln4494del, tendo o estudo familiar revelado que a co-segregação com a hipercolesterolemia na família não era completa (Figura III.2.10A). Por se tratar de 1 alteração que apresenta uma deleção *in frame*, apenas foi possível utilizar os programas SIFT e Mutation Taster para a análise *in silico*. O programa SIFT classificou-a como patogénica e o Mutation Taster como polimorfismo. A alteração foi pesquisada, mas não detetada num painel de 95 indivíduos portugueses normolipidémicos nem nem na base de SNP (<http://www.ncbi.nlm.nih.gov/snp>) ou *National Heart, Lung, and Blood Institute* (NHLBI) *database* (<http://evs.gs.washington.edu/EVS/>) (Tabela III.2.5). A alteração foi pesquisada, mas não detetada, num painel de 95 indivíduos com diagnóstico clínico de FH mas sem alteração num dos três genes associados à FH. O estudo funcional a partir das partículas de LDL do CI mostrou existir uma diminuição de cerca de 44% na ligação e internalização quando comparado com o controlo normal (Figura III. 2.10B e C).

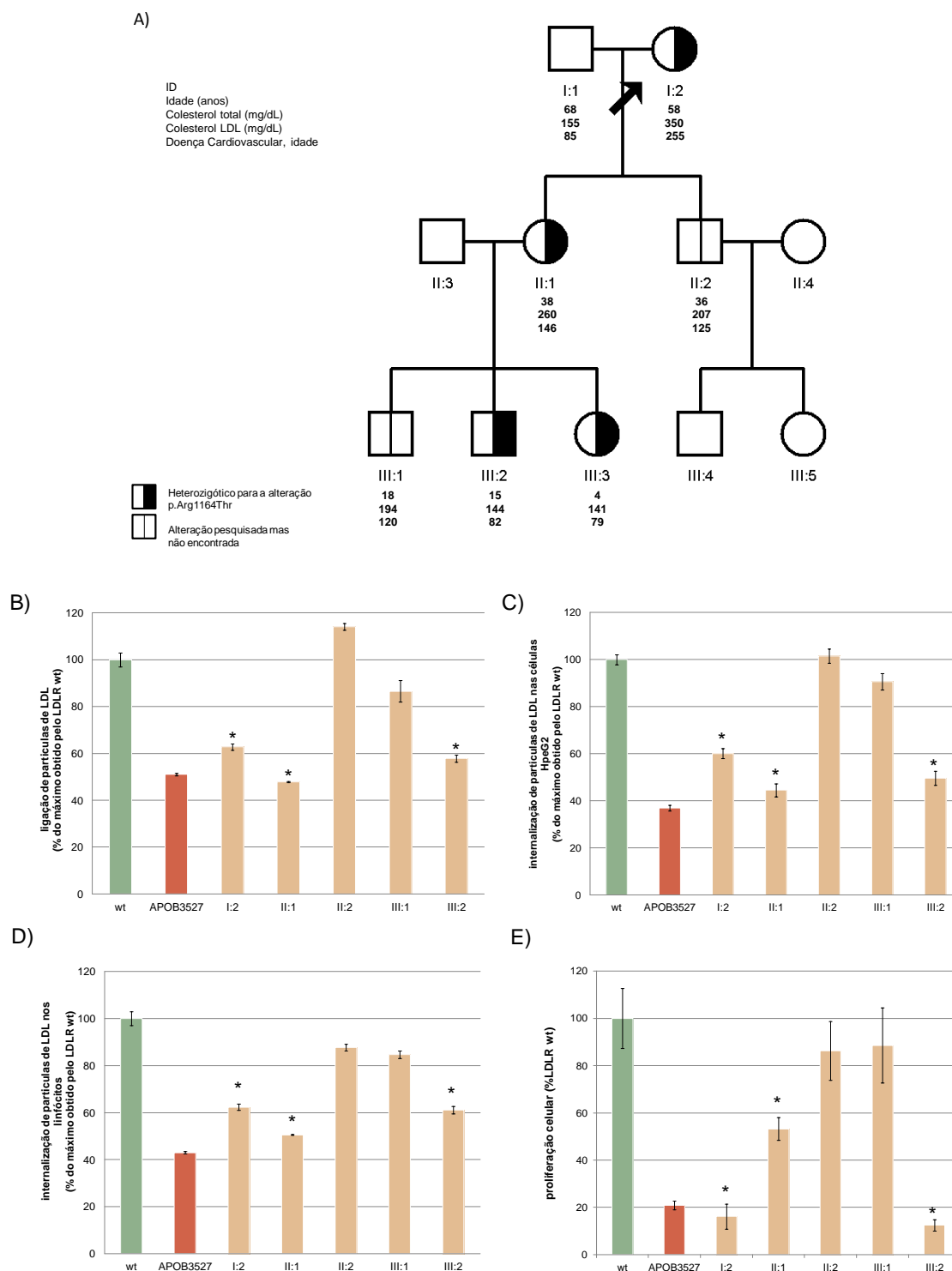


Figura III.2.9 – Resultados do estudo funcional do CI 25023. A) Árvore genealógica do CI 25023. A seta representa o caso índice. B) Ensaio de ligação do complexo LDLR:LDL em linfócitos com partículas de LDL marcada após uma incubação de 4 horas a 4°C. C) Ensaio de internalização das partículas de LDL com marcação FITC em linfócitos wt após uma incubação de 4 horas a 37°C LDL D) Níveis de internalização das partículas de LDL marcada em células HepG2 após uma incubação de 4 horas a 37°C. E) Proliferação celular das células U937 após uma incubação de 48 horas, o cálculo da proliferação celular foi realizado por comparação dos indivíduos com as alterações no gene APOB vs controlo normal (sem adição de LDL). Os dados apresentados representam a média de pelo menos 3 determinações \*  $p < 0,05$  entre os indivíduos com a alteração p.Arg1164Thr e LDLR wt (teste t-sudent).

A mãe do CI apresenta uma hipercolesterolemia mais severa e o estudo funcional demonstrou uma redução da ligação do complexo LDLR:LDL bem como da internalização do complexo semelhante ao controle APOB3527. O mesmo resultado foi observado nas células HepG2 e U937 (Figura III. 2.10D e E). No entanto a avó materna do caso índice apresenta valores elevados de colesterol e a alteração não foi encontrada. Por outro lado, o avô e o tio materno do CI apresentam uma hipercolesterolemia ligeira e são heterozigóticos para a alteração p.Gln4494del. O estudo funcional de ambos os indivíduos em linfócitos e linhas celulares de HepG2 mostrou existir uma redução na ligação e internalização de cerca de 40% enquanto para a avó os dados mostram uma ligação e internalização normais (Figura III. 2.10B, C, D e E).

Também nesta família foi analisada a presença de alterações nos genes PCSK9 e APOC3 que originam um fenótipo de hipocolesterolemia, não tendo sido detetada qualquer alteração. No CI 28120 foi realizado um estudo preliminar de *exome sequencing* de forma a identificar a causa da hipercolesterolemia na família, uma vez que não tinha sido realizado até ao momento o estudo funcional da alteração encontrada no gene APOB. O estudo do exoma no CI 28120 identificou a alteração p.Asp19His no gene *ATP-binding cassette subfamily G member 8 (ABCG8)* no CI. O estudo da co-segregação da alteração na família verificou que a mãe, tio e avó materna apresentavam também a alteração no gene ABCG8 (Figura III.2.10A).

Os ensaios funcionais para as alterações p.Asp1113His e p.Tyr1247Cys, nas diferentes três linhas celulares, confirmaram que as mesmas são polimorfismos uma vez que não afetam a ligação da partícula de LDL ao LDLR nem a sua internalização (Figura III.2.11).

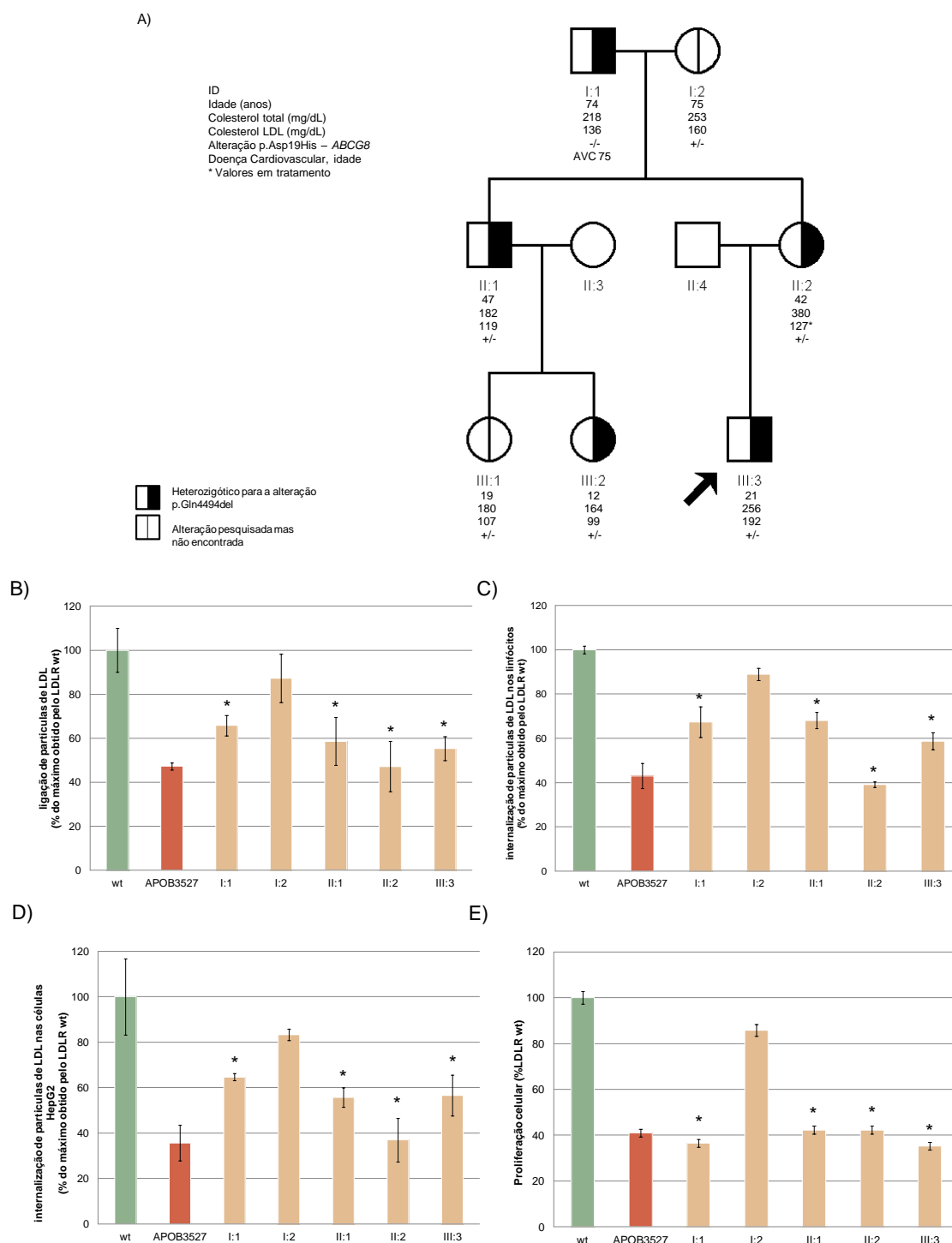


Figura III.2.10 – Resultados do estudo funcional do CI 28120. A) Árvore genealógica do CI 28120. A seta representa o caso índice. AVC, acidente vascular cerebral. Esta família apresenta também uma alteração no gene ABCG8, indicada na Figura como +/- e -/-; B) Ensaio de ligação do complexo LDLR:LDL em linfócitos com partículas de LDL marcada por FITC após uma incubação de 4 horas a 4°C. C) Ensaio de internalização das partículas de LDL com marcação FITC em linfócitos wt após uma incubação de 4 horas a 37°C LDL D) Níveis de internalização das partículas de LDL marcada com FITC em células HepG2. E) Proliferação celular das células U937 após uma incubação de 48 horas, o cálculo da proliferação celular foi realizado por comparação dos indivíduos com as alterações no gene APOB vs controle normal (sem adição de LDL). Os dados apresentados representam a média de pelo menos 3 determinações. \*  $p < 0,05$  entre as os indivíduos com a alteração p.Gln4494del e LDLR wt (teste t-student).

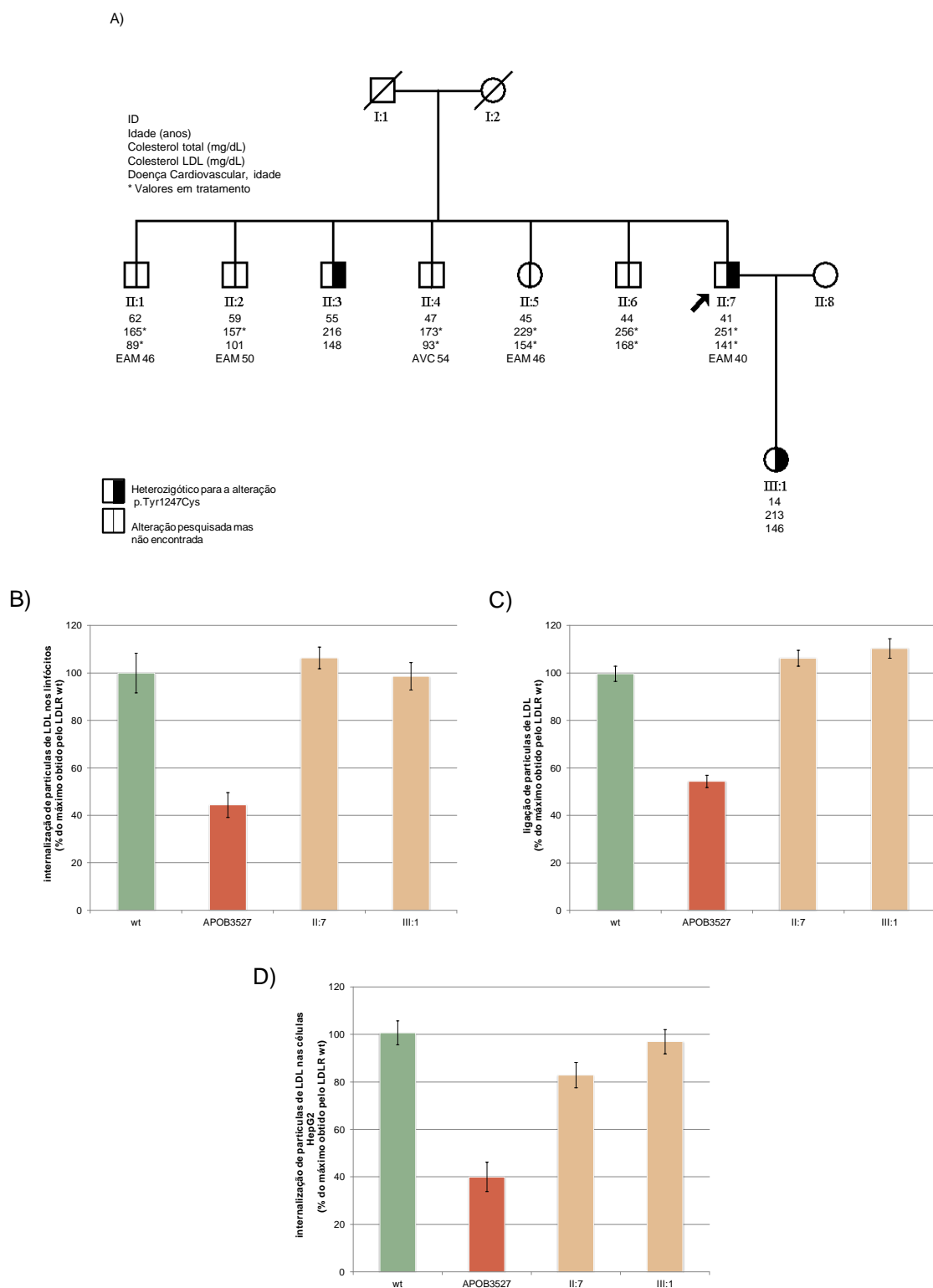


Figura III.2.11 – Resultados do estudo funcional do CI 26086. A) Árvore genealógica do CI 26086. A seta representa o caso índice. AVC, acidente vascular cerebral. EAM, enfarte agudo do miocárdio; B) Ensaio de ligação do complexo LDLR:LDL em linfócitos com partículas de LDL marcada após uma incubação de 4 horas a 4°C; C) Ensaio de internalização das partículas de LDL com marcação FITC em linfócitos wt após uma incubação de 4 horas a 37°C LDL; D) Níveis de internalização das partículas de LDL marcada em células HepG2 após uma incubação de 4 horas a 37°C. \*  $p < 0,05$  entre as os indivíduos com a alteração p.Thr1247Cys e LDLR wt (teste t-sudent).

Tabela III.2.3 – Caracterização, frequência, análise *in silico*, co-segregação e classificação funcional das alterações encontradas no gene *APOB*.

Exon	Alteração		Co-segregação	Frequência			in silico				Estudo funcional		
	cDNA	Proteína		Painel Normolip.	1000G	ESP	SIFT	PolyPhen2	Condel	Mutation Taster		Grantham Score	phyloP
19	c.2981C>T	p.Pro994Leu	1/2*/0/0	0/96	0,10%	0,10%	Patogénica	Provável	Patogénica	Causa doença.	Mod. Conserv.	4	Não realizado
22	c.3337G>C	p.Asp1113His	1/2;2/2	2/96	0,30%	0,80%	Patogénica	Possível	Patogénica	Polimorfismo	Mod. Conserv.	0,93	Neutra
22	c.3491G>C	p.Arg1164Thr	2/2;2/4	0/96	nd	nd	Patogénica	Possível	Patogénica	Polimorfismo	Mod. Conserv.	0,69	Patogénica
24	c.3740A>G	p.Tyr1247Cys	2/7;1/1	0/96	0,10%	0,05%	Patogénica	Provável	Patogénica	Causa doença.	Radical	0,85	Neutra
26	c.5599C>T	p.Arg1867Trp	2/2;0/0	0/96	nd	0,02%	Patogénica	Provável	Patogénica	Causa doença.	Mod. Radical.	0,21	Não realizado
26	c.5690G>A	p.Arg1897His	3/4;0/1	0/96	nd	nd	Neutra	Neutra	Neutra	Polimorfismo	Conservado	np	Não realizado
26	c.6639_6641delTGA	p.Asp2213del	?	1/96	nd	0,32%	Patogénica	np	np	Causa doença.	np	0,29	Não realizado
26	c.7853T>C	p.Ile2618Thr	?	0/96	nd	nd	Neutra	Neutra	Neutra	Polimorfismo	Mod. Conserv.	1,74	Não realizado
26	c.9835A>G	p.Ser3279Gly	2/2;0/1	2/96	0,70%	0,80%	Patogénica	Provável	Neutra	Polimorfismo	Mod. Conserv.	1,74	Não realizado
29	c.13480_13482delCAG	p.Gln4494del	2/3;3/4	0/96	nd	nd	Patogénica	np	np	Polimorfismo	np	4,45	Patogénica

Normolip, normolipídemicos; ESP, exome sequencing project; Mod. Conserv., moderate conservative; Co-segregação: com alteração/com fenótipo; com alteração/sem fenótipo; \* os dois indivíduos possuem a mutação p.Ala431Thr no LDLR; dbSNP – base dados SNP; nd – não descrito; np – não é possível a previsão.

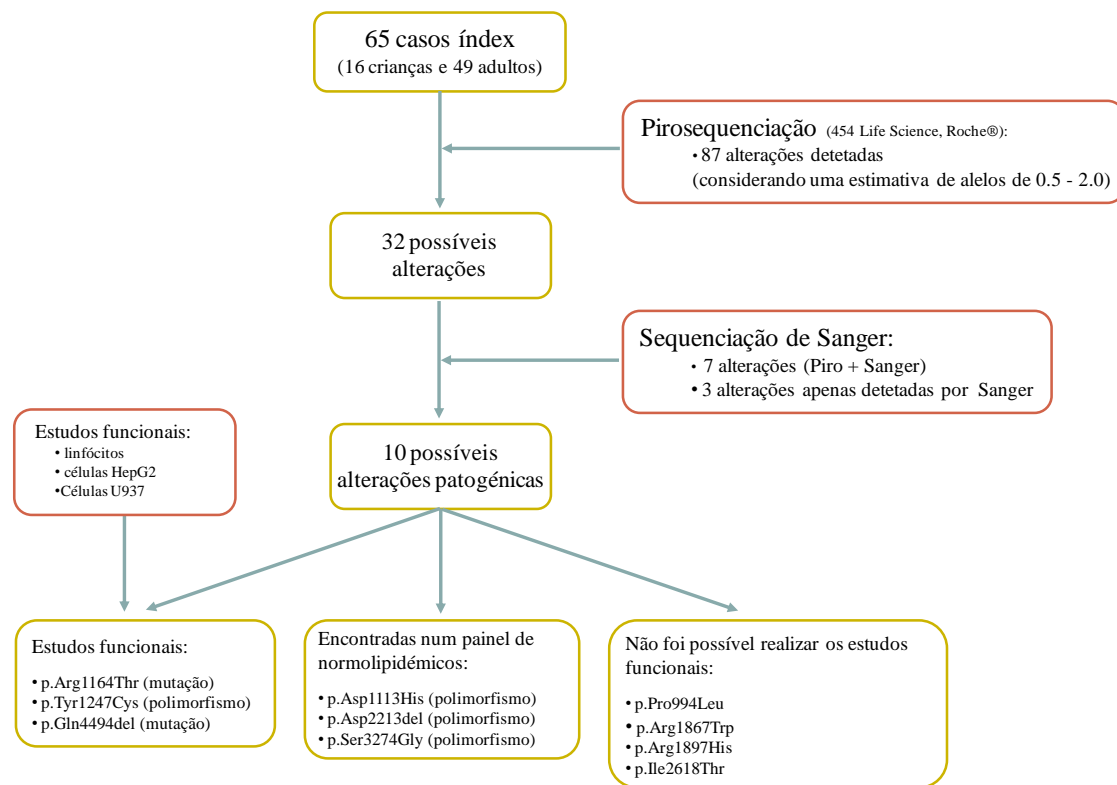


Figura II:2.12 – Diagrama resumo da identificação e caracterização de alterações encontradas no gene *APOB*.



### III.2.6 Classificação das mutações

Como referido anteriormente, apesar de terem sido identificadas 117 alterações no gene *LDLR*, apenas 45 têm estudos funcionais comprovando a sua patogenicidade, significando que dos 294 CI identificados com 1 alteração apenas 192 destes apresentam uma mutação comprovadamente patogénica. Se apenas considerarmos estes CI como tendo 1 alteração num dos três genes associados à FH a taxa de deteção tornar-se-ia muito baixa (30%) e poderíamos estar a classificar (erradamente) alguns CI como não tendo uma mutação causadora de doença. De modo a classificar as alterações sem estudos funcionais, foi necessário desenvolver um critério baseado em programas bioinformáticos que realizam uma previsão do efeito da alteração na proteína. Assim e de modo a realizar uma validação da análise *in silico* foram analisadas através de programas bioinformáticos 27 mutações *missense* e 17 mutações de *splicing*, identificadas no gene *LDLR* com estudos funcionais (Tabelas III.2.4 e III.2.5).

Cada alteração foi classificada como provavelmente patogénica, possivelmente patogénica ou neutra, conforme descrito no ponto II.6 do capítulo II. A classificação de provavelmente patogénica foi atribuída considerando os seguintes critérios: provavelmente patogénica ou possivelmente patogénica para o programa *Polyphen2*; patogénica para o programa *SIFT*; causa doença para o programa *Mutation Taster*; moderadamente radical ou radical para o *Grantham score* e um valor superior a 4,4 para o programa *PhyloP*. Com base nestes pressupostos as alterações *missense* sem estudos funcionais foram classificadas da seguinte forma:

- Alterações prováveis – uma alteração foi considerada provável quando 5 ou 6 dos programas bioinformáticos utilizados a classificaram como uma alteração patogénica.
- Alterações possíveis – uma alteração foi considerada possível quando somente 3 ou 4 dos programas bioinformáticos a classificaram como uma alteração patogénica.
- Neutra – uma alteração foi considerada neutra sempre que pelo menos 4 programas bioinformáticos a classificaram como tal.

Este critério classificou corretamente 22/27 alterações *missense* (81,5%) através da análise *in silico*. Após a validação verificou-se que a melhor combinação apresentava uma sensibilidade de 81,5% e especificidade de 100%. Todos os programas falharam

pelo menos uma vez a previsão *in silico* a uma alteração, mas o programa com pior resultado na previsão *in silico* foi o *Grantham score*, classificando corretamente apenas 12 das 27 alterações analisadas (Tabela II.2.4).

Os critérios estabelecidos para as alterações de *splicing* encontram-se descritos no ponto II.6 do capítulo II, as alterações foram consideradas patogénicas quando a previsão dos 3 programas originava um resultado inferior a 80%. Os critérios utilizados para a classificação *in silico* das alterações de *splicing* sem estudos funcionais, foram:

- Alterações prováveis – uma alteração foi considerada provável quando pelo menos 2 programas bioinformáticos a classificaram como patogénica (*score* inferior a 80%).
- Alteração neutra – uma alteração foi considerada neutra quando pelo menos 2 programas bioinformáticos a classificaram com *score* superior a 80%.

Com base neste pressupostos classificou-se corretamente 17 alterações (100%) através da análise *in silico* (Tabela III.2.5). Para as alterações de *splicing* todos os programas falharam pelo menos uma vez na previsão de uma alteração. No entanto, para as alterações de *splicing* obteve-se uma sensibilidade e especificidade de 100%.

De acordo com estes critérios classificaram-se 36 alterações encontradas no EPHF no gene *LDLR*, para as quais não foram realizados estudos funcionais (Tabelas III.2.6 e III.2.7), tendo sido obtido os seguintes resultados: 19 prováveis, 7 possíveis e 10 neutras.

Os CI com alterações classificadas como possíveis ou neutras foram considerados como não sendo FH negativos e analisados como tal nas seguintes etapas.

Tabela III.2.4 – Análise *in vitro* vs *in silico* das alterações *missense* encontradas no gene *LDLR*

Alteração		Co-segreg.		Painel Normolip.	Análise in silico					Previsão in silico *		Referência
c.DNA	Proteína				Polyphen2	SIFT	CONDEL	Mutation Taster	Grantham score	PhyloP		
c.226G>T	p.Gly76Trp	3/3; 5/6	0/95	Provável	Patogénica	Neutra	Causa Doença	Radical	4,16	Possível	Neutra	(Bourbon et al. 2008)
c.265T>C	p.Cys89Arg	3/3; 0/1	0/95	Provável	Patogénica	Patogénica	Causa Doença	Radical	4,81	Provável	Patogénica	(Bertolini et al. 1999)
c.326G>T	p.Cys109Phe	4/4; 0/0	0/95	Possível	Patogénica	Patogénica	Causa Doença	Radical	4,64	Provável	Patogénica	(Hobbs et al. 1992)
c.502G>C	p.Asp168His	2/2; 0/3	0/95	Provável	Patogénica	Patogénica	Causa Doença	Mod. Conserv.	5,86	Provável	Patogénica	(Leitersdorf et al. 1993)
c.530C>T	p.Ser177Leu	30/31; 0/5	0/95	Provável	Patogénica	Patogénica	Causa Doença	Mod. Radical	4	Provável	Patogénica	(Hobbs et al. 1988)
c.590G>T	p.Cys197Phe	3/3; 0/2	0/95	Provável	Patogénica	Patogénica	Causa Doença	Radical	3,76	Provável	Patogénica	(Hobbs et al. 1992)
c.662A>G	p.Asp221Gly	2/2; 0/0	0/95	Provável	Patogénica	Patogénica	Causa Doença	Mod. Conserv	4,73	Provável	Patogénica	(Hobbs et al. 1992)
c.670G>A	p.Asp224Asn	43/44; 0/10	0/95	Provável	Neutra	Patogénica	Causa Doença	Conservative	5,86	Possível	Patogénica	(Hobbs et al. 1992)
c.682G>A	p.Glu228Lys	2/2; 0/0	0/95	Provável	Patogénica	Patogénica	Causa Doença	Mod. Conserv	5,86	Provável	Patogénica	(Leitersdorf et al. 1993)
c.796G>A	p.Asp266Asn	1/1;0/0	0/95	Provável	Patogénica	Patogénica	Causa Doença	Mod. Conserv	5,61	Provável	Patogénica	(Fouchier et al. 2001)
c.806G>A	p.Gly269Asp	8/10; 1/9	0/95	Neutra	Neutra	Neutra	Causa Doença	Mod. Conserv.	2,87	Neutra	Neutra	(Etxebarria et al. 2012)
c.829G>A	p.Glu277Lys	1/1; 1/2	0/95	Neutra	Neutra	Neutra	Causa Doença	Mod. Conserv	3,03	Neutra	Neutra	(Pereira et al. 1995)
c.1027G>A	p.Gly343Ser	2/3; 0/3	0/95	Provável	Patogénica	Patogénica	Causa Doença	Mod. Conserv.	4,08	Possível	Patogénica	(Hobbs et al. 1992)
c.1222G>A	p.Glu408Lys	3/3; 0/0	0/95	Provável	Patogénica	Patogénica	Causa Doença	Mod. Conserv.	5,53	Provável	Patogénica	(Hobbs et al. 1992)
c.1285G>C	p.Val429Leu	1/1; 0/0	0/95	Neutra	Neutra	Neutra	Causa Doença	Conservative	3,84	Neutra	Patogénica	(Silva et al. 2012)
c.1291G>A	p.Ala431Thr	79/83; 0/14	0/95	Provável	Patogénica	Patogénica	Causa Doença	Mod.conserv.	5,53	Provável	Patogénica	(Hobbs et al. 1990)

Tabela III.2.4 – Continuação

Alteração		Co-segreg.	Painel Normolip.	Análise <i>in silico</i>				Previsão <i>in silico</i> *		Estudo funcional	Referência
c.DNA	Proteína			Polyphen2	SIFT	CONDEL	Mutation Taster	Grantham score	PhyloP		
c.1322T>C	p.Ile441Thr	9/10; 0/2	0/95	Possível	Patogénica	Patogénica	Causa Doença	Mod. Conserv.	4,56	Provável	Patogénica (Fouchier et al. 2001) **
c.1432G>A	p.Gly478Arg	5/6; 1/2	0/95	Provável	Patogénica	Patogénica	Causa Doença	Mod. Radical	5,45	Provável	Patogénica (Hobbs et al. 1992)
c.1468T>C	p.Trp490Arg	4/4; 0/1	0/95	Provável	Patogénica	Patogénica	Causa Doença	Mod. Radical	4,48	Provável	Patogénica (Bourbon et al. 2008)
c.1633G>T	p.Gly545Trp	33/34; 0/4	0/95	Provável	Patogénica	Patogénica	Causa Doença	Radical	1,66	Provável	Patogénica (Bourbon et al. 2008) **
c.1690A>C	p.Asn564His	1/10/0	0/95	Provável	Patogénica	Patogénica	Causa Doença	Mod. radical	4,89	Provável	Patogénica (Jensen et al. 1996)
c.1775G>A	p.Gly592Glu	19/21; 0/0	0/95	Provável	Patogénica	Patogénica	Causa Doença	Mod. radical	5,77	Provável	Patogénica (Hobbs et al. 1992)
c.1942T>C	p.Ser648Pro	1/1; 0/0	0/95	Neutra	Patogénica	Neutra	Polimorfismo	Mod.conserv.	0,93	Neutra	Patogénica (Bourbon et al. 2008)
c.2054C>T	p.Pro685Leu	8/8; 0/1	0/95	Provável	Patogénica	Patogénica	Causa Doença	Mod.conserv.	5,69	Provável	Patogénica (Soutar et al. 1989)
c.2093G>T	p.Cys698Phe	4/5; 0/0	0/95	Provável	Patogénica	Patogénica	Causa Doença	Radical	5,69	Provável	Patogénica Nova **
c.2177C>T	p.Thr726Ile	12/13; 3/4	0/95	Neutra	Patogénica	Neutra	Polimorfismo	Mod.conserv	1,01	Neutra	Neutra (Hobbs et al. 1992)
c.2575G>A	p.Val859Met	2/3; 1/2	0/95	Neutra	Patogénica	Patogénica	Polimorfismo	Conservativa	-0,76	Neutra	Neutra (Bourbon et al. 2008)

Co-segregação: com alteração/com fenótipo; com alteração/sem fenótipo; Normolip, normolipídénicos; Cau Doença: causa doença; Polimorf, polimorfismo; Mod. Conserv., moderadamente conservado; \* Previsão *in silico* com base em 6 programas bioinformáticos considerados. A sublinhado encontram-se as alterações cuja análise *in vitro* e *in silico* dão iguais. \*\*Estudo funcional descrito pela primeira neste trabalho.

Tabela III.2.5 – Análise *in vitro* vs *in silico* das alterações de *splicing* encontradas no gene *LDLR*

Alteração		Co-segregação	Painel Normolip.	Análise <i>in silico</i>			Previsão <i>in silico</i> *	Estudo funcional	Referência
Intrão	c.DNA			Splice Port	NNSP	NetGen2			
2	c.190+4insTG	4/5; 0/0	0/95	0 (exão 2D)	0 (exão 2D)	0 (exão 2D)	Provável	Patogénica	(Medeiros et al. 2010)
3	c.313+1G>A	1/1; 0/0	0/95	0 (exão 3D)	0 (exão 3D)	0 (exão 3D)	Provável	Patogénica	(Leren et al. 1994)
3	c.313+1G>T	2/2; 0/0	0/95	0 (exão 3D)	0 (exão 3D)	0 (exão 3D)	Provável	Patogénica	(Lombardi et al. 1995)
3	c.313+1G>C	1/1; 0/0	0/95	0 (exão 3D)	0 (exão 3D)	0 (exão 3D)	Provável	Patogénica	(Lombardi et al. 1995)
3	c.313+6T>C	2/2; 0/0	0/95	46 (exão 3D)	45 (exão 3D)	75 (exão 3D)	Possível	Patogénica	(Bourbon et al. 2009)
5	c.818-3C>G	3/3; 0/1	0/95	0 (exão 6A)	0 (exão 6A)	0 (exão 6A)	Provável	Patogénica	Nova
5	c.818-2A>G	9/10; 0/4	0/95	0 (exão 6A)	0 (exão 6A)	0 (exão 6A)	Provável	Patogénica	(Medeiros et al. 2010)
7	c.1060+1G>A	14/15; 0/3	0/95	0 (exão 7D)	0 (exão 7D)	0 (exão 7D)	Provável	Patogénica	(Medeiros et al. 2010)
7	c.1061-8T>C	12/13; 3/4	2/75	100 (exão 8A)	100 (exão 8A)	100 (exão 8A)	Neutra	Neutra	(Bourbon et al. 2009)
9	c.1359-5C>G	5/7; 1/1	0/75	74 (exão 10A)	74 (exão 10A)	99 (exão 10A)	Provável	Patogénica	(Bourbon et al. 2009)
11	c.1706-10G>A	3/4; 1/1	0/95	100 (exão 12A)	100 (exão 12A)	100 (exão 12A)	Neutra	Neutra	(Cenarro et al., 1998)
12	c.1846-1G>A	1/1; 0/0	0/95	0 (exão 13A)	0 (exão 13A)	0 (exão 13A)	Provável	Patogénica	(Hobbs et al. 1992)
14	c.2140+1G>A	3/3; 0/1	0/75	0 (exão 14D)	0 (exão 14D)	0 (exão 14D)	Provável	Patogénica	(Takada et al. 2002)
14	c.2140+5G>A	4/4; 0/0	0/75	86 (exão 14D)	33 (exão 14D)	86 (exão 14D)	Neutra	Neutra	(Bourbon et al. 2009)
17	c.2389G>T (p.Val797Leu)	3/3; 0/1	0/75	44 (exão 16D)	82 (exão 16D)	60 (exão 16D)	Provável	Patogénica	(Mak et al. 1998)
17	c.2389G>A (p.Val797Met)	6/6; 0/0	0/75	43 (exão 16D)	53 (exão 16D)	60 (exão 16D)	Provável	Patogénica	(Bourbon et al. 2009)
17	c.2547+1G>A	6/7; 0/1	0/75	0 (exão 17D)	0 (exão 17D)	0 (exão 17D)	Provável	Patogénica	(Medeiros et al. 2010)

Co-segregação: com alteração/com fenótipo; com alteração/sem fenótipo; Normolip. nomolipidémicos. \* Previsão *in silico* com base nos 3 programas bioinformáticos considerados. \*\*Estudo funcional descrito pela primeira neste trabalho.

Tabela III.2.6 – Classificação das alterações *missense* encontradas no *gene LDLR* sem estudos funcionais

Alteração		Análise <i>in silico</i>							Previsão		Referência
c.DNA	Proteína	Co-segregação	Painel	Polyphen2	SIFT	CONDEL	Mutation Taster	Grantham score	phyloP	<i>in silico</i> *	
c.58G>A	p.Gly20Arg	1/2; 0/0	0/95	Neutra	Neutra	Neuro	Polimorfismo	Mod. radical	0,12	Neutra	(Anselmen et al. 2002)
c.241C>T	p.Arg81Cys	2/2; 0/0	0/95	Provável	Patogénica	Patogénica	Causa Doença	Radical	3,76	Provável	(Nissen et al. 1998)
c.274C>G	p.Gln92Glu	1/1; 0/0	0/95	Neutra	Neutra	Neuro	Polimorfismo	Conservativa	0,93	Neutra	(Medeiros et al. 2010)
c.292G>C	p.Gly98Arg	1/1; 0/0	0/95	Possível	Patogénica	Patogénica	Causa Doença	Mod. Conserv.	2,38	Possível	(Bourbon et al. 2008)
c.326G>T	p.Cys109Phe	3/3; 0/0	0/95	Possível	Patogénica	Patogénica	Causa Doença	Radical	5,69	Provável	(Hobbs et al. 1992)
c.427T>C	p.Cys143Arg	3/3; 0/0	0/95	Possível	Patogénica	Patogénica	Causa Doença	Radical	4,73	Provável	(Nauck et al. 2001)
c.473C>T	p.Ser158Cys	1/1; 0/0	0/95	Possível	Patogénica	Patogénica	Causa Doença	Mod. Radical	1,42	Provável	(Medeiros et al. 2010)
c.551G>A	p.Cys184Tyr	11/11; 0/2	0/95	Possível	Patogénica	Patogénica	Causa Doença	Radical	5,37	Provável	(Lee et al. 1998)
c.631C>G	p.His211Asp	4/4; 0/1	0/95	Possível	Patogénica	Patogénica	Causa Doença	Mod. Conserv.	5,86	Provável	(Bourbon et al. 2008)
c.661G>T	p.Asp221Tyr	5/5; 0/1	0/95	Possível	Patogénica	Patogénica	Causa Doença	Radical	5,86	Provável	(Koivisto et al. 1995)
c.666C>G	p.Cys222Trp	1/1; 0/0	0/95	Possível	Patogénica	Patogénica	Causa Doença	Radical	3,11	Provável	<b>Nova</b>
c.862G>A	p.Glu288Lys	2/2; 1/1	0/95	Provável	Neutra	Patogénica	Causa Doença	Mod. Conserv.	5,69	Possível	(Varret et al. 1998)
c.1088C>A	p.Thr363Asn	1/1; 1/1	0/95	Possível	Patogénica	Patogénica	Causa Doença	Mod. Conserv.	0,53	Possível	<b>Nova</b>
c.1216C>T	p.Arg406Trp	35/36; 0/2	0/95	Possível	Patogénica	Patogénica	Causa Doença	Mod. radical	2,55	Possível	(Reshef et al. 1996)
c.1325A>G	p.Tyr442Cys	1/1; 0/0	0/95	Provável	Patogénica	Patogénica	Causa Doença	Radical	2,87	Provável	(Wang et al. 2005)
c.1352T>C	p.Ile451Thr	2/2; 0/0	0/95	Possível	Patogénica	Patogénica	Causa Doença	Mod. Conserv.	4,56	Provável	(Miltiadous et al. 2000)
c.1403T>A	p.Val468Asp	1/1; 0/0	0/95	Provável	Patogénica	Patogénica	Causa Doença	Radical	4,24	Provável	(Medeiros et al. 2010)
c.1417A>G	p.Ile473Val	3/3; 0/0	0/95	Neutra	Neutra	Neuro	Polimorf	Conservativa	0,21	Neutra	<b>Nova</b>
c.1455C>G	p.His485Gln	2/2; 1/2	0/95	Provável	Patogénica	Neuro	Causa Doença	Conservativa	-0,28	Possível	(Bourbon et al. 2008)
c.1463T>C	p.Ile488Thr	4/4; 0/0	0/95	Provável	Patogénica	Patogénica	Causa Doença	Mod. Conserv.	4,48	Provável	(Fouchier et al. 2001)
c.1802A>T	p.Asp601Val	5/9; 0/2	0/95	Provável	Patogénica	Patogénica	Causa Doença	Radical	2,06	Provável	<b>Nova</b>
c.1816G>T	p.Ala606Ser	6/7; 0/0	0/95	Neutra	Neutra	Neuro	Causa Doença	Mod. Conserv.	1,34	Neutra	(Sun et al. 1997)
c.1840T>A	p.Phe614Ile	2/2; 0/1	0/95	Provável	Patogénica	Patogénica	Causa Doença	Conservativa	4,32	Possível	(Bourbon et al. 2008)
c.1876G>A	p.Glu626Lys	8/9; 0/1	0/95	Possível	Patogénica	Neuro	Causa Doença	Mod. Conserv.	1,81	Neutra	(Fouchier et al. 2005)

Tabela III.2.6 – Continuação

Alteração		Co-segregação		Painel Normolip.		Análise <i>in silico</i>						Previsão <i>in silico</i> *	Referência
c.DNA	Proteína					Polyphe2	SIFT	CONDEL	Mutation Taster	Grantham score	phyloP		
c.1960C>T	p.Leu654Phe	1/1; 0/0	0/95	0/95	0/95	Provável	Patogénica	Patogénica	Causa Doença	Conservativa	2,14	Possível	Nova
c.1966C>A	p.His656Asn	4/7; 1/2	0/95	0/95	0/95	Possível	Patogénica	Patogénica	Causa Doença	Mod. Conserv.	5,69	Provável	(Mozas et al. 2004)
c.1999T>C	p.Cys667Ser	1/1; 0/0	0/95	0/95	0/95	Provável	Patogénica	Patogénica	Causa Doença	Radical	4,73	Provável	(Vergopoulos et al.)
c.2146G>A	p.Glu716Lys	1/1 0/0	0/95	0/95	0/95	Neutra	Neutra	Neuro	Polimorfismo	Mod. Conserv.	0,93	Neutra	(Medeiros et al. 2010)
c.2291T>C	p.Ile764Thr	3/3; 0/1	0/95	0/95	0/95	Neutra	Neutra	Neuro	Polimorfismo	Conservativa	-0,2	Neutra	Nova

Co-segregação: com alteração/com fenótipo; com alteração/sem fenótipo; Normolip, normolipídicos; Mod. Conserv., moderadamente conservado; \* Previsão *in silico* com base em 6 programas bioinformáticos considerados. A sublinhado encontram-se as alterações classificadas como prováveis

Tabela III.2.7 – Classificação das alterações de *splicing* e sinónimas encontradas no *gene LDLR* sem estudos funcionais

Alteration		Co-segreg.	Protein	Normolip.		Análise <i>in silico</i>			Previsão <i>in silico</i> *	Referência
cDNA				Painel	Splice Port	NNSSP	NetGen2	Mutation Taster		
c.1586+2T>A		1/1;0/0		0/95	0 (exão 10D)	0 (exão 10D)	0 (exão 10D)	np	Provável	Nova
c.1845+1delG		1/1;0/0		0/95	0 (exão 12D)	0 (exão 12D)	0 (exão 12D)	np	Provável	(Bourbon et al. 2009)
c.90C>T	p.Asn30Asn	1/1;0/0		0/95	100 (exão 2A) 100 (exão2D)	99 (exão 2A) 54 (exão2D)	100 (exão 2A) 100 (exão 2D)	Causa Doença	Neutra	(Lind et al. 2002)
c.1185G>C	p.Val395Val	6/6;0/1		0/95	0 (exão 8 A&D)	0 (exão 8 A&D)	0 (exão 8 A&D)	Disease Causing	Provável	(Bourbon et al. 2009)
c.1279A>C	p.Arg427Arg	6/7;1/1		0/95	100 (exão 9A) 100 (exão 9D)	100 (exão 9A) 100 (exão 9D)	100 (exão 9A) 60 (exão 9D)	Causa Doença	Neutra	Nova
c.1383C>T	p.Gly461Gly	2/2;0/0		0/95	100 (exão 10A) 100 (exão 10D)	100 (exão 10A) 100 (exão 10D)	100 (exão 10A) 100 (exão 10D)	Polymorphism	Neutra	Nova
c.1875C>T	p.Asn625Asn	3/3;0/1		0/95	100 (exão 13A) 100 (exão 13D)	100 (exão 13A) 83 (exão13D)	100 (exão 13A) 100 (exão 13D)	Causa Doença	Neutra	(Widhalm et al. 2007)

Co-segregação: com alteração/com fenótipo; com alteração/sem fenótipo; Normolip. normolipídemicos. \* Previsão *in silico* com base em 3 programas bioinformáticos considerados. A sublinhado encontram-se as alterações classificadas como prováveis







Parte 3 – Caracterização Bioquímica de doentes com  
diagnóstico clínico de FH:  
genótipo vs fenótipo

Este trabalho está incluído nos seguintes artigos:

**Alves AC, et al.** (2014) submitted  
Medeiros AM, **Alves AC, et al.** (2014) *J Lip Res* *acepated*



### **Parte 3 – Caracterização Bioquímica de doentes com diagnóstico clínico de FH –genótipo vs fenótipo**

Durante os anos de 1999 a 2012, o Estudo Português de Hipercolesterolemia Familiar recebeu 642 casos índice (250 crianças e 392 adultos) com diagnóstico clínico de FH e 1461 familiares. Após o estudo molecular foram identificados com uma alteração provavelmente causadora de FH 114 CI do grupo pediátrico e 180 CI do grupo adultos. No entanto, alguns dos CI apresentavam uma alteração cuja patogenicidade não estava comprovada e foi necessário realizar estudos funcionais de forma a classificar a alteração, como descrito na parte 2 deste capítulo. As alterações não analisadas do ponto de vista funcional foram classificadas tendo em conta um critério de classificação baseado em 6 programas bioinformáticos que catalogou as alterações em prováveis, possíveis e neutras, como explicado anteriormente na parte 2 deste capítulo.

Para a caracterização bioquímica dos indivíduos em estudo neste trabalho foram considerados como tendo uma alteração patogénica todos os indivíduos que apresentassem alterações classificadas como provavelmente patogénicas e as alterações com estudos funcionais. Assim, foram considerados para análise 84 CI do grupo pediátrico e 127 CI do grupo adultos. A identificação molecular destes CI permitiu ainda identificar mais 296 familiares com FH (75 crianças e 221 adultos).

A análise estatística foi realizada tendo em conta dois grupos distintos, o grupo pediátrico (dos 2 aos 17 anos) e o grupo adultos. De forma a avaliar-se o perfil bioquímico de todos os indivíduos referenciados ao EPHF dividiram-se os CI com base no diagnóstico molecular: CI com uma alteração definitiva (diagnóstico definitivo - DD), CI com uma alteração provavelmente patogénica (diagnóstico provável - DP), CI com uma alteração possivelmente patogénica (diagnóstico provável - DPos), CI com uma alteração neutra (diagnóstico neutro - DN) e CI sem alteração (FH negativo – FH<sup>-</sup>). Para a análise de todos os parâmetros bioquímicos utilizaram-se os valores obtidos antes do tratamento hipolipemiante. Os 8 CI homozigóticos e heterozigóticos compostos foram retirados da análise estatística geral e serão apresentados separadamente, uma vez que o fenótipo da forma homozigótica difere substancialmente da forma heterozigótica.

### III.3.1 Grupo Pediátrico

Foram referenciados ao EPHF 642 CI, dado que, 250 pertenciam ao grupo pediátrico, destes 141 eram do sexo feminino e 109 do sexo masculino. A idade média era de  $10,1 \pm 3,6$  anos. Todos os CI referenciados apresentavam hipercolesterolemia, uma vez que 46,9% das crianças cumpriam apenas dieta controlada e 33,1% efetuavam, também, terapêutica farmacológica. Em nenhuma das crianças se observou a presença de xantomas tendinosos. Verificou-se que 39,3% das crianças eram obesas ou tinham excesso de peso e que 66,7% praticavam menos de 3 horas por semana de atividade física. A percentagem de CI crianças com familiares de primeiro grau com DCV era de 24,5% (Tabela III.3.1).

Tabela III.3.1. – Caracterização clínica do grupo pediátrico, observada no momento de referência ao EPHF

n=250	
<b>Dados demográficos</b>	
Idade (anos)	$10,1 \pm 3,6$
Género masculino	43,6%
<b>Medidas terapêuticas</b>	
Dieta	46,9%
Medicação + dieta	33,1%
<b>História Familiar</b>	
Familiar de 1º ou 2º grau com DCV	24,5%
<b>Outras patologias</b>	
Hipertensão (pre-hipertensão e hipertensão)	35,3%
<b>Estilo de vida</b>	
Exercício físico (> 3h/semana)	33,3%
Hábitos tabágicos	0%
<b>IMC</b>	
Normal	63,5 %
Excesso de peso	16,5 %
Obesidade	14,1 %

DCV, doença cardiovascular; Medicação: a tomar estatinas e/ou ezetimiba e/ou resinas e/ou fibratos; Exercício físico referido pelo CI; h, horas.

### III.3.1.1 Caracterização Bioquímica vs diagnóstico molecular

De forma a se avaliar o perfil bioquímico do grupo pediátrico os CI foram divididos com base no diagnóstico molecular: DD = 71, DP=13; DPos =13; DN =17 e FH<sup>-</sup>=135.

Verificou-se que o valor de colesterol total encontrava-se mais elevado nas crianças com DD e DP ( $315,8 \pm 56,7$  e  $308,9 \pm 62,0$  mg/dL, respetivamente) em comparação com as crianças que apresentavam DPos, DN ou FH<sup>-</sup> ( $249,8 \pm 29,9$ ;  $271,7 \pm 62,9$  e  $250,6 \pm 37,6$  mg/dL, respetivamente), verificando-se a existência de diferenças estatisticamente significativas entre os CI com DD e DPos, DD e DN, DD e FH<sup>-</sup> ( $p < 0,001$ ). No que diz respeito aos valores de cLDL verificou-se que existiam diferenças estatisticamente significativas quando comparadas as crianças com DD de todos os outros diagnósticos ( $246,2 \pm 55,3$ ,  $182,8 \pm 26,8$ ;  $193,7 \pm 49,9$  e  $175,4 \pm 34,6$  mg/dL, respetivamente), incluindo o DP ( $205,9 \pm 44,1$  mg/dL) mas enquanto a diferença estatística entre DP, DPos, DN e FH<sup>-</sup> é bastante significativa ( $p > 0,001$ ) no caso do DP vs DD esta é ligeiramente significativa ( $p = 0,04$ ). No entanto os valores de apoB entre as crianças com DD e DP não apresentam diferenças estatisticamente significativas ( $124,2 \pm 25,1$ ,  $135,5 \pm 20,3$  mg/dL, respetivamente) enquanto entre as crianças com DD e os restantes diagnósticos apresentam diferenças estatisticamente significativas (Figura III.3.1) (Tabela III.3.2). Para o cHDL e triglicéridos não se verificaram diferenças estatisticamente significativas entre os grupos (tabela III.3.2).

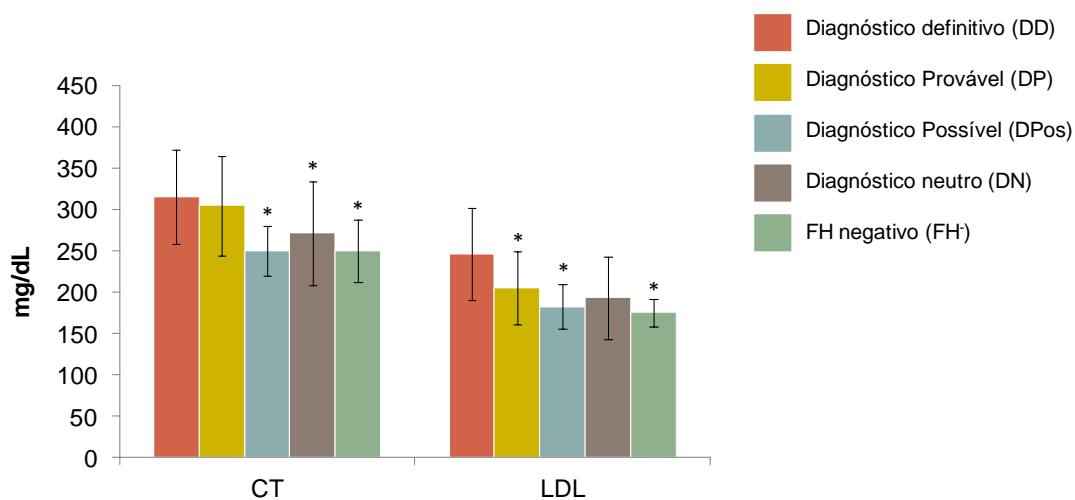


Figura III.3.1 – Níveis de colesterol total e cLDL do grupo pediátrico por diagnóstico molecular. \*Diferença estatisticamente significativa quando comparados indivíduos com DD vs DPos, DN e FH<sup>-</sup> para o CT e FH diagnóstico definitivo vs restantes diagnósticos para o cLDL (T-Test ou teste Mann-Whitney).

Tabela III.3.2 – Caracterização bioquímica dos CI do grupo pediátrico de acordo com a classificação molecular

	Diagnóstico Definitivo (n=71)	Diagnóstico Provável (n=13)	Diagnóstico Possível (n=13)	Diagnóstico Neutro (n=17)	FH negativo (n=136)
<b>Perfil lipídico básico (mg/dl)<sup>a</sup></b>	<b>n=70</b>	<b>n=13</b>	<b>n=13</b>	<b>n=16</b>	<b>n=136</b>
Colesterol Total	315,8±56,7	308,9±62,0	249,8±29,9*	271,7±62,9*	250,6±37,6*
cLDL	246,2±55,3	205,9±44,1*	182,8±26,8*	193,7±49,9*	175,4±34,6*
cHDL	51,1±12,8	57,6±21,2	56,1±14,6	63,4±18,0	59,9±16,2
Triglicerídeos	89,4±43,9	103,1±91,7	75,8±32,1	74,2±27,6	99,4±49,5
não-cHDL:cHDL razão	5,6±1,9	4,6±1,8	3,7±1,2*	3,6±1,1*	3,4±1,3*
<b>Perfil lipídico específico (mg/dl)<sup>b</sup></b>	<b>n=39</b>	<b>n=7</b>	<b>n=9</b>	<b>n=12</b>	<b>n=98</b>
Lp(a)	45,7±45,6	24,4±24,4	54,3±47,5	95,5±125,8	58,5±69,8
apoA-I	135,1±24,9	141,3±22,1	146,3±21,5	152,0±23,5	155,9±29,9
apoB	124,2±25,1	135,5±20,3	97,3±15,9*	112,5±80,9*	89,6±16,5*
apoB: apoA-I razão	0,95±0,29	0,89±0,23	0,68±0,19*	0,73±0,40*	0,59±0,19*
<b>Perfil lipídico diferenciado (mg/dl)<sup>c</sup></b>	<b>n=26</b>	<b>n=6</b>	<b>n=4</b>	<b>n=5</b>	<b>n=65</b>
apoA-II	26,0±5,1	26,5±4,3	28,2±1,4	32,7±6,5	29,7±5,9
apoC-II	3,0±1,2	3,8±1,9	2,9±0,9	4,3±1,3	4,0±1,6
apoC-III	7,4±1,8	7,8±1,4	6,9±1,2	7,7±1,1	8,9±2,5
apoE	3,4±1,0	3,4±0,7	3,3±1,1	4,0±1,1	2,9±0,8†
sdLDL	43,4±17,0	43,6±14,3	27,0±10,8*	25,4±11,8*	25,5±11,1*

Resultados expressos em média ± desvio padrão (DP); FH: Hipercolesterolemia Familiar; <sup>a</sup> valores mais altos reportados antes de tratamento; <sup>b</sup> perfil realizado no INSA no momento de referência ao estudo dos CI sem tratamento; <sup>c</sup> perfil realizado no INSA, apenas, a partir de 2010, valores antes de tratamento. \*p<0,05 entre CI com FH definitiva comparado individualmente com cada classificação molecular (T-Test ou teste Mann-Whitney); †p<0,05 entre CI FH negativo comparado individualmente com cada categoria molecular (T-Test).

No que respeita à avaliação do perfil lipídico específico, verificou-se que nos valores de de apolipoproteína A-I (apoA-I) e razão apoB:apoA-I as crianças com DD e DP apresentam uma razão superior aos CI com os restantes diagnósticos (DPos e DN) e sem mutação, sendo a diferença estatisticamente significativa (p<0,001) (Tabela III.3.2).

A avaliação do perfil específico das lipoproteínas apolipoproteína A-II (apoA-II), apolipoproteína C-II (apoC-II) e apolipoproteína C-III (apoC-III) revelou não existirem diferenças estatisticamente significativas entre os CI com os diferentes diagnósticos moleculares. No entanto, para as sdLDL verificou-se existirem diferenças estatisticamente significativas entre os CI com diagnóstico definitivo (43,4±17,0 mg/dL) e os CI com diagnóstico possível, neutro e FH negativo (27,0±10,8; 25,4±11,8 e 25,5±11,1 mg/dL) (p<0,001) tal como verificado para os valores de cLDL (Tabela III.3.2). A comparação dos valores de apoE entre os CI com diferentes diagnósticos



moleculares revelou existirem diferenças estatisticamente significativas ( $p<0,05$ ) entre os CI sem mutação e as restantes classificações, dado que as crianças sem mutação são as que apresentam o valor mais baixo de apoE e curiosamente as crianças com diagnóstico molecular neutro as que apresentam o maior valor (Tabela III.3.2).

#### III.3.1.2 Caracterização Bioquímica grupo pediátrico - tipo de mutações

De forma a caracterizar bioquimicamente o grupo pediátrico quanto ao tipo de mutação (mutações nos genes *LDLR*, *APOB* e *PCSK9*) considerou-se para análise todos os CI e familiares com diagnóstico molecular definitivo e provável de FH pertencentes a este grupo. Desta forma, foram consideradas 159 crianças com DD e DP (85 CI e 74 familiares) com uma mutação num dos três genes associados à FH, uma vez que 85 destas crianças tinham uma mutação pontual, 54 uma mutação nula, 16 uma mutação de *splicing* e 3 uma mutação no gene *APOB*. Apesar das crianças com mutações nulas apresentarem valores mais elevados de colesterol total e cLDL, verificou-se que apenas existiam diferenças estatisticamente significativas ( $p=0,012$ ) para os valores de CT entre as crianças com alterações no gene *APOB* quando comparadas com as crianças com mutações nulas no gene *PCSK9* ( $251,0\pm31,6$  e  $313,4\pm60,3$  respetivamente). Não se verificam diferenças estatisticamente significativas nos restantes parâmetros bioquímicos para estas alterações (Figura III.3.2) (Tabela III.3.3). No entanto verificou-se existirem diferenças estatisticamente significativas ( $p<0,05$ ) nos valores de apoB ( $145,1\pm30,6$  e  $116,7\pm7,2$  mg/dL respetivamente) e consequentemente razão apoB:apoA-I ( $1,2\pm0,2$  e  $0,87\pm0,1$  mg/dL respetivamente) nas crianças com alterações de *splicing* quando comparadas com as crianças com alterações no gene *APOB*.

O perfil bioquímico diferenciado foi iniciado apenas em 2010. Para se comparar os parâmetros bioquímicos associados a este perfil era necessário analisar somente as crianças que não cumpriam tratamento farmacológico e que apresentavam uma mutação vs sem mutação. Verificou-se não existir tamanho amostral para esta análise optando-se por não realizar tratamento estatístico para o perfil diferenciado.

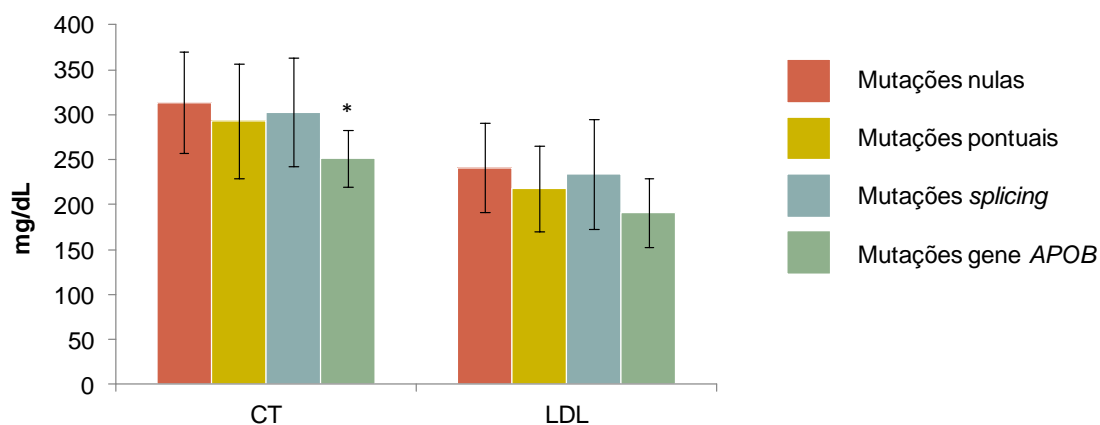


Figura III.3.2 – Níveis de colesterol total e cLDL do grupo pediátrico vs tipo de mutação.

Tabela III.3.3 – Caracterização bioquímica dos doentes do grupo pediátrico de acordo com o tipo de mutação

	Gene <i>LDLR</i>			Gene <i>APOB</i>
	Mutações nulas (n=54)	Mutações pontuais (n=85)	Mutações <i>splicing</i> (n=16)	
<b>Perfil lipídico básico (mg/dl)<sup>a</sup></b>	<b>n=52</b>	<b>n=79</b>	<b>n=14</b>	<b>n=4</b>
Colesterol Total	313,4±60,3*	292,9±56,2	303,1±64,2	251,0±31,6*
cLDL	241,3±61,1	217,4±49,5	233,9±47,6	191,0±38,2
cHDL	52,1±11,9	52,8±14,7	44,5±10,9	49,0±4,0
Triglicerídeos	99,0±35,9	82,2±35,9	99,4±55,8	43,0±28,6
não-cHDL:cHDL razão	5,3±1,8	4,9±1,7	6,2±2,4	4,0±0,4
<b>Perfil lipídico específico (mg/dl)<sup>b</sup></b>	<b>n=32</b>	<b>n=37</b>	<b>n=8</b>	<b>n=4</b>
Lp(a)	45,4±48,4	49,8±46,2	43,5±46,0	35,3±28,6
apoA-I	131,5±23,3	128,8±25,3	124,0±19,6	133,3±8,1
apoB	126,6±38,4	121,9±26,6	145,1±30,6*	116,7±7,2*
apoB: apoA-I razão	1,0±0,4	1,0±0,5	1,2±0,2*	0,87±0,1*

Resultados expressos em média ± desvio padrão (DP); <sup>a</sup> valores mais altos reportados antes de tratamento; <sup>b</sup> perfil realizado no INSA no momento referência ao estudo dos CI sem tratamento. \*p<0,05 entre as mutações respetivamente assinaladas.

### III.3.1.3 Caracterização Bioquímica – casos índice vs familiares

Encontra-se descrito que familiares de doentes com FH e com mutação num dos genes associados à FH, apresentação um fenótipo mais moderado que os CI (Sun *et al.* 1998; Talmud *et al.* 2013). Foi realizada uma comparação dos parâmetros bioquímicos do perfil lipídico básico e avançado de forma a averiguar se o mesmo sucedia na amostra em estudo. Para esta análise foram considerados apenas os doentes que apresentavam um diagnóstico molecular definitivo ou provável. Verificou-se que existiam diferenças estatisticamente significativas entre CI e familiares com mutação num dos genes associados à FH para o colesterol total, cLDL, apoA-I e razão não-cHDL:cHDL (Tabela III.3.4). Os familiares apresentam uma diminuição de cerca de 10% (32 mg/dL) nos valores de CT quando comparados os CI. Verificou-se existir, também, uma diminuição de cerca de 12% (28mg/dL) nos valores de cLDL dos familiares, em comparação com os CI.

### III.3.1.4 Análise de outros fatores de risco cardiovascular: crianças com FH vs crianças sem FH

Realizou-se uma comparação do número de fatores de riscos nas crianças com uma alteração definitiva ou provável vs sem uma alteração de forma a verificar quanto factores de risco para além da hipercolesterolemia estas crianças já apresentavam. Verificou-se que cerca de 40% das crianças com FH apresentam já outros 2 fatores de risco e que apenas existem diferenças estatisticamente significativas nas crianças com 4 fatores de risco em ambos os grupos ( $p=0.008$ ) (Figura III.3.3).

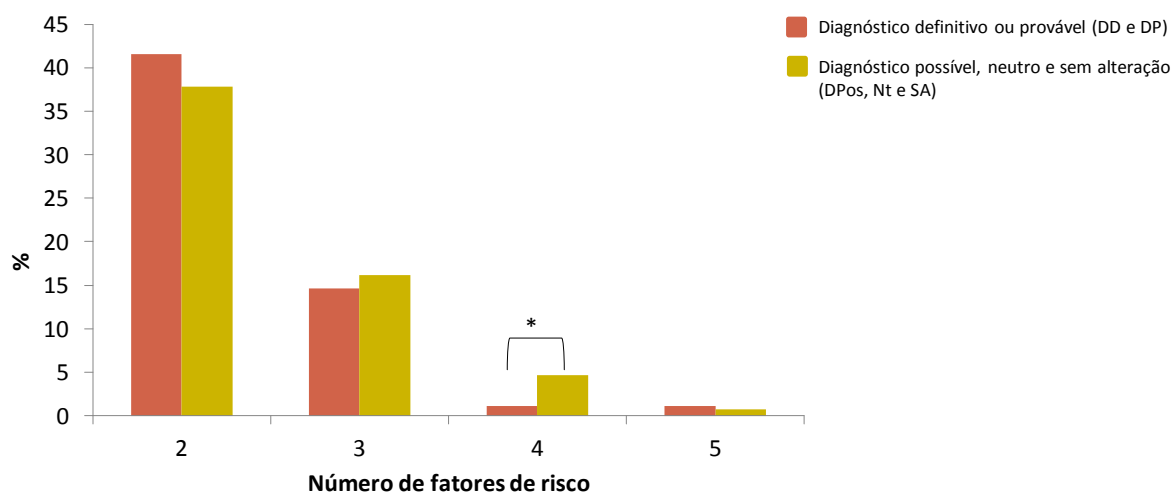


Figura III.3.3 – Número de fatores de risco no grupo pediátrico. \* Diferença estatisticamente significativa entre os grupos para a presença de 4 fatores de risco (Teste  $\chi^2$ ).

### III.3.2 Grupo Adultos

Foram referenciados ao EPHF 389 CI pertencentes ao grupo adultos, sendo 216 do sexo feminino e 173 do sexo masculino, com uma idade média de  $43,9 \pm 13,9$  anos. Todos os CI referenciados apresentavam hipercolesterolemia (CT superior a 290 mg/dL), dado que 7,3% dos CI cumpriam apenas dieta e 71,9% efetuavam, também, terapêutica farmacológica. Observou-se que 2,3% dos CI adultos apresentavam xantomas tendinosos, uma característica clínica de FH definitiva. Verificou-se que 55,9% dos CI adultos eram obesos ou apresentavam excesso de peso; e apenas 23,6% praticava atividade física. A percentagem de CI com DCV nesta amostra foi de 19,3% (Tabela III.3.5).

#### III.3.2.1 Caracterização Bioquímica grupo adultos – diagnóstico molecular

Para o grupo adulto, procedeu-se ao mesmo tipo de análise estatística e por conseguinte à mesma separação efetuada no grupo das crianças para a caracterização bioquímica, tendo em consideração os diferentes diagnósticos moleculares: DD (n=98), DP (n=21), DPos (n=31), DN (n=19) e FH (n=212) (Tabela III.3.5).

Verificou-se que o valor de colesterol total é mais elevado nos adultos com DD e DP ( $378,7 \pm 80,2$  e  $375,1 \pm 59,0$  mg/dL) em comparação com os doentes com DPos, DN ou FH ( $327,0 \pm 46,8$ ;  $328,0 \pm 43,3$  e  $319,6 \pm 75,4$  mg/dL) sendo estas diferenças

estatisticamente significativas para os CI com DD e DPos; DD e DN, DD e FH<sup>-</sup> (p<0,001).

Tabela III.3.4 – Comparação bioquímica de casos index vs familiares do grupo pediátrico com FH

	<b>Caso índice</b> <i>n</i> =85	<b>Familiar</b> <i>n</i> =75
<b><i>Perfil lipídico básico (mg/dl)<sup>a</sup></i></b>	<b><i>n</i>=84</b>	<b><i>n</i>=66</b>
Colesterol Total*	314,1±57,1	281,6±56,2
cLDL*	239,8±55,5	211,8±49,6
cHDL	52,2±14,6	50,9±11,9
Triglicéridos	91,51±53,5	85,1±540,8
não-cHDL:cHDL razão *	5,4±1,9	4,77±1,7
<b><i>Perfil lipídico específico (mg/dl)<sup>b</sup></i></b>	<b><i>n</i>=46</b>	<b><i>n</i>=46</b>
Lp(a)	42,4±43,5	52,1±48,4
apoA-I*	136,0±24,4	123,0±21,1
apoB	124,0±24,2	126,37,9
apoB: apoA-I razão	0,9±0,3	1,1±0,6

Resultados expressos em média ± desvio padrão (DP); <sup>a</sup> valores mais elevados reportados antes de tratamento; <sup>b</sup> perfil realizado no INSA no momento de referência ao estudo, CI sem tratamento. \*p<0.05 entre casos índice e familiares (T-Test ou teste Mann-Whitney).

O mesmo foi observado para os valores de cLDL para os CI com diagnóstico definitivo e provável, 292,3±83,5 e 280,7±83,5 mg/dL, respetivamente; em comparação com os adultos com diagnóstico possível, neutro e sem mutação: 241,1±47,1; 229,7±39,0 e 217,4±48,2 mg/dL (Tabela III.3.6) (Figura III.3.4) sendo estas diferenças estatisticamente significativas quando comparados os CI com DD e DPos; DD e DN, DD e FH<sup>-</sup> (p<0,001). Verificou-se, igualmente, para a razão não-cHDL:cHDL existir um ligeiro aumento nos CI com diagnóstico molecular definitivo sendo este estatisticamente significativo (p<0,001). Para os valores de cHDL verificou-se existirem diferenças estatisticamente significativas entre os CI com diagnóstico molecular definitivo vs neutro, sendo os valores de cHDL mais elevados nos CI com diagnóstico neutro. Para os triglicéridos não se observaram diferenças estatisticamente significativas entre os grupos (Tabela III.3.6).

Tabela III.3.5 – Caracterização clínica do grupo adultos, observada no momento de referenciação ao EPHF

Dados demográficos		Outras patologias	
Idade (anos)	43,9±13,9	Hipertensão	10,7%
Género masculino	44,4%	Diabetes	1,7%
Estilo de vida		IMC	
Exercício físico	23,6%	Normal	42,4%
Hábitos tabágicos	9,8%	Excesso de peso	38,0%
Hábitos etanólicos	12,6%	Obesidade	17,9%
Doença Cardiovascular		Sintomas físicos	
Sim	19,3%	Xantomas tendinosos	2,3%
Medidas terapêuticas			
Dieta	7,3%		
Medicação	71,9%		

Medicação: a tomar estatinas e/ou ezetimiba e/ou resinas e/ou fibratos e/ou LDL aferese; Exercício físico referido

No que respeita a avaliação do perfil bioquímico específico, verificou-se que os valores de apoB são mais elevados nos adultos com diagnóstico molecular definitivo, provável e possível ( $172,9 \pm 64,7$ ;  $135,5 \pm 64,7$  e  $143,2 \pm 52,8$  mg/dL, respetivamente), em comparação com os CI com diagnóstico molecular neutro ou FH negativo ( $112,6 \pm 51,4$  e  $109,5 \pm 37,9$  e mg/dL, respetivamente) sendo estas diferenças estatisticamente significativas quando comparados os CI com DD e DN, DD e FH<sup>-</sup> ( $p < 0,001$ ). Para a razão apoB:apoA-I verificou-se, também, que esta é superior nos CI com alterações DD e DP em comparação com os CI com diagnóstico molecular possível e neutro. Para os valores bioquímicos de apoA-I e Lp(a) não foram encontradas diferenças estatisticamente significativas (Tabela III.3.6).

A avaliação do perfil bioquímico diferenciado das lipoproteínas apoA-II, apoC-II e apoC-III apenas foi realizado para os CI com diagnóstico molecular definitivo, provável e sem mutação, uma vez que apenas existia 1 CI com valores antes de tratamento para o diagnóstico possível e neutro. Não foram detetadas quaisquer diferenças estatisticamente significativas entre os CI com diferente diagnósticos moleculares. No entanto, para as sdLDL verificou-se existirem diferenças estatisticamente significativas entre os CI com diagnóstico definitivo e provável ( $57,1 \pm 38,9$  e  $60,8 \pm 54,5$  mg/dL,

$p < 0,05$ , respetivamente) e os CI sem mutação ( $40,6 \pm 20,1$  mg/dL) sendo estas diferenças estatisticamente significativas ( $p < 0,001$ ) (tabela III.3.6).

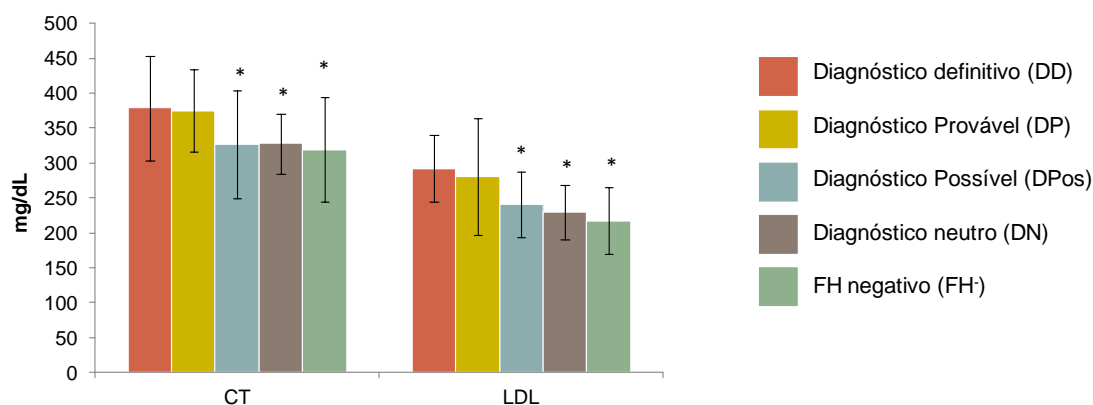


Figura III.3.4 – Níveis de colesterol total e cLDL do grupo adultos por diagnóstico molecular.

\* Diferença estatisticamente significativa quando comparado CI com diagnóstico definitivo vs possível, neutro e FH (T-Test ou teste Mann-Whitney).

Tabela III.3.6 – Caracterização bioquímica dos CI do grupo adultos de acordo com a classificação molecular

	Diagnóstico Definitivo (n=98)	Diagnóstico Provável (n=21)	Diagnóstico Possível (n=31)	Diagnóstico Neutro (n=19)	FH negativo (n=215)
<b>Perfil lipídico básico (mg/dl) <sup>a</sup></b>					
	n=88	n=17	n=30	n=15	n=188
Colesterol Total	378,7±80,2	375,1±59,0	327,0±76,8*	328,0±43,3*	319,6±75,4*
cLDL	292,3±83,5	280,7±83,5	241,1±47,1*	229,7±39,0*	217,4±48,2*
cHDL	48,7±14,2	59,2 ±19,8	53,69±14,4	59,8±9,9*	58,4±17,3*
Triglicerídeos	144,8±66,2	123,9±64,4	151,3±61,2	160,9±67,9	159,5±70,6
não-cHDL:cHDL razão	7,2±2,9	6,0±1,9	5,3±1,7*	4,7±1,3*	4,9±2,2
<b>Perfil lipídico avançado (mg/dl) <sup>b</sup></b>					
	n=13	n=5	n=6	n=3	N=36
Lp(a)	55,2±34,9	58,0±34,6	21,3±15,3	15,1±13,1	65,6±56,7
apoA-I	157,8±36,9	142,6±45,9	167,7±47,2	179,3±12,0	167,0±52,3
apoB	172,9±64,7	135,5±64,7	143,2±52,8	112,6±51,4*	109,5±37,9*
apoB: apoA-I razão	1,3±0,8	1,1±0,4	0,8±0,4*	0,6±0,3*	1,0±2,2*
<b>Perfil lipídico específico (mg/dl) <sup>c</sup></b>					
	n=7	n=4	n=1	n=1	n=20
apoA-II	33,5±7,5	24,7±3,5	---	---	31,7±6,2
apoC-II	4,6±1,5	3,8±2,5	---	---	4,7±2,4
apoC-III	11,8±3,9	9,4±2,5	---	---	11,9±3,9
apoE	3,9±1,7	3,1±1,1	---	---	3,2±0,9
sdLDL	57,1±38,9	60,8±54,5	---	---	40,6±20,1*

Resultados expressos em média ± desvio padrão (DP); FH: Hipercolesterolemia Familiar; <sup>a</sup> valores mais elevados reportados antes de tratamento; <sup>b</sup> perfil realizado no INSA no momento referencição ao estudo dos CI sem tratamento; <sup>c</sup> perfil realizado no INSA, apenas, a partir de 2010, os valores são antes de

tratamento. \* $p < 0,05$  entre CI com FH definitiva comparado individualmente por cada classificação molecular (T-Test ou teste Mann-Whitney).

### III.3.2.2 Caracterização Bioquímica grupo adulto- diagnóstico molecular

De forma a caracterizar bioquimicamente os doentes do grupo adulto com alterações nos genes *LDLR*, *APOB* e *PCSK9* considerou-se para esta análise todos os CI e familiares com diagnóstico molecular definitivo e provável de FH que pertenciam ao grupo adultos.

Foram identificados 342 adultos (122 CI e 220 familiares) com uma mutação num dos três genes associados à FH, dado que, 185 apresentavam uma mutação pontual, 94 uma mutação nula, 40 uma mutação de *splicing* no gene *LDLR*. No gene *APOB* foram identificados 18 CI e 5 CI uma alteração no gene *PCSK9*. Verificou-se que os doentes com mutações nulas apresentam valores de CT e cLDL ( $373,7 \pm 97,1$  e  $285,7 \pm 92,0$  mg/dL, respetivamente) mais elevados em comparação com os doentes com mutações pontuais ( $340,6 \pm 73,4$  e  $252,0 \pm 66,9$  mg/dL, respetivamente) e no gene *APOB* ( $313,1 \pm 58,3$  e  $229,3 \pm 47,2$  mg/dL, respetivamente), sendo esta diferença estatisticamente significativa ( $p < 0,05$ ) quando comparados os CI com alterações nulas vs alterações pontuais e no gene *APOB* (Tabela III.3.7) (Figura III.3.5).

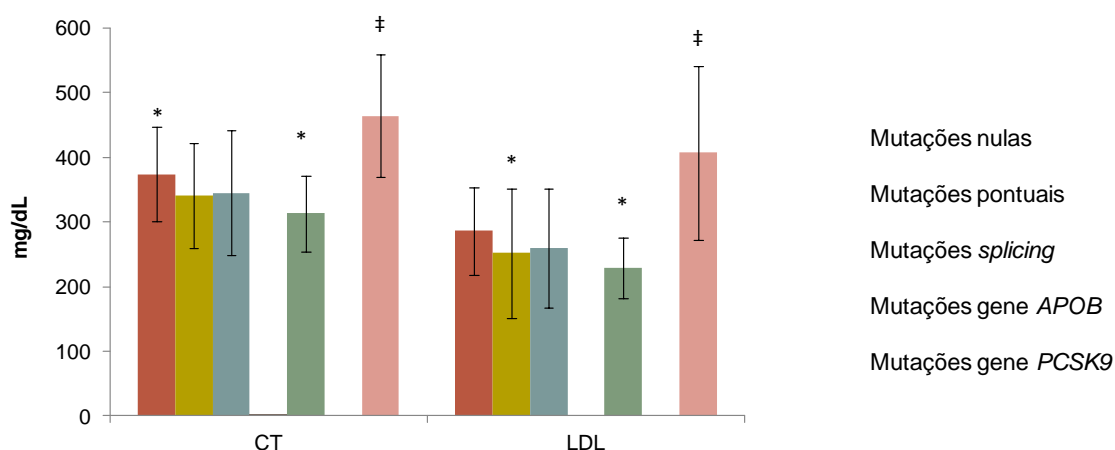


Figura III.3.5 – Níveis de colesterol total e cLDL do grupo adultos tipo de mutação.

\*Diferença estatisticamente significativa quando comparado CI com mutações nulas vs mutações pontuais e mutações no gene *APOB* (T-Test ou teste Mann-Whitney). ‡ Diferença estatisticamente significativa quando comparado CI com alterações no *PCSK9* vs restantes mutações (T-Test ou teste Mann-Whitney).



No que diz respeito aos valores bioquímicos dos doentes com mutações de *splicing*, não se verificou a existência de diferenças estatisticamente significativas quando comparados com doentes com mutações nulas, pontuais e no gene *APOB* (Tabela III.3.7).

Verificou-se que os doentes identificados no EPHF com uma mutação no gene *PCSK9* apresentavam um fenótipo mais severo que os restantes doentes, sendo as diferenças estatisticamente significativas para todos os parâmetros bioquímicos analisados ( $p<0,001$ ) (Tabela III.3.7 e Figura III.3.5).

Uma vez que o perfil lipídico específico foi iniciado apenas em 2010, o tamanho amostral de doentes adultos com diagnóstico molecular definitivo ou provável, sem medicação e com parâmetros bioquímicos doseados era reduzido, optando-se por não realizar a análise destes parâmetros para os diferentes tipos de mutação.

Tabela III.3.7 – Caracterização bioquímica dos doentes do grupo adultos de acordo com o tipo de mutação

	Gene <i>LDLR</i>			Gene <i>APOB</i> (n=17)	Gene <i>PCSK9</i> (n=3)
	Mutações nulas (n=68)	Mutações pontuais (n=144)	Mutações <i>splicing</i> (n=35)		
<b>Perfil lipídico básico (mg/dl)<sup>a</sup></b>	<b>n=56</b>	<b>n=110</b>	<b>n=23</b>	<b>n=15</b>	<b>n=3</b>
Colesterol Total	373,7±97,1	340,6±73,4*	345,8±81,9	313,1±58,3*	463,9±94,5*
cLDL	285,7±92,0	252,0±66,9*	259,3±100,0	229,3±47,2*	406,9±134,3*
cHDL	51,5±16,4	51,3±12,3	46,6±15,5	55,1±16,2	46,1±5,2
Triglicerídeos	144,6±67,9	141,5±70,8	105,4±42,3	106,7±61,3	169,3±67,7
não-cHDL:cHDL razão	6,8±3,0	5,9±2,2	6,8±3,2	5,0±1,6	9,5±4,0
<b>Perfil lipídico avançado (mg/dl)<sup>b</sup></b>	<b>n=26</b>	<b>n=65</b>	<b>n=12</b>	<b>n=7</b>	<b>n=0</b>
Lp(a)	47,4±40,4	46,9±41,9	37,7±31,9	47,8±32,0	---
apoA-I	144,3±36,1	141,4±27,8	139,9±31,4	132,1±19,4	---
apoB	160,5±48,4	158,5±50,9	147,4±42,6	132,4±28,9	---
apoB: apoA-I razão	1,2±0,4	1,2±0,5	1,1±0,4	1,0±0,2	---

Resultados expressos em média ± desvio padrão (DP); <sup>a</sup> valores mais elevados reportados antes de tratamento; <sup>b</sup> perfil realizado no INSA no momento da referência ao estudo dos CI, sem tratamento; \* $p<0,05$  mutações nulas comparadas individualmente com restantes mutações (T-Test ou teste Mann-Whitney).

### III.3.2.3 Caracterização Bioquímica – casos índice vs familiares

Para o grupo dos adultos foi, também, comparado o fenótipo entre os casos índice e os familiares com mutação. Para a análise foram considerados apenas os doentes que apresentavam um diagnóstico molecular definitivo ou provável. Verificou-se, que tal

como no grupo pediátrico, existiam diferenças estatisticamente significativas entre CI e familiares com mutação num dos genes associados à FH para o colesterol total, cLDL, apoA-I e razão não-cHDL:cHDL (Tabela III.3.8).

Verificou-se existirem diferenças nos níveis de CT entre CI e familiares, uma vez que os CI apresentavam um valor de CT 13% superior (47,6 mg/dL) quando comparado com o valor de CT dos familiares. Também os níveis de cLDL são superiores, em cerca de 17% (49,3 mg/dL) nos CI quando comparado com os familiares.

Tabela III.3.8 – Comparação bioquímica de casos index vs familiares do grupo adultos com FH

	<b>Caso índice</b> <i>n=120</i>	<b>Familiar</b> <i>n=220</i>
<b><i>Perfil lipídico básico (mg/dl)<sup>a</sup></i></b>	<b><i>n=105</i></b>	<b><i>n=162</i></b>
Colesterol Total*	378,1±76,9	330,5±81,0
cLDL*	290,6±78,9	241,3±75,2
cHDL	50,7±15,9	51,1±13,1
Triglicerídeos	141,9±66,0	132,2±69,2
não-cHDL:cHDL razão *	7,0±2,8	5,7±2,4
<b><i>Perfil lipídico específico (mg/dl)<sup>b</sup></i></b>	<b><i>n=46</i></b>	<b><i>n=46</i></b>
Lp(a)	55,9±33,7	44,4±40,5
apoA-I	153,6±38,8	138,6±27,1
apoB	164,1±59,9	154,6±46,2
apoB: apoA-I razão	1,2±0,7	1,1±0,4

Resultados expressos em média ± desvio padrão (DP); <sup>a</sup> valores mais altos reportados antes de tratamento; <sup>b</sup> perfil realizado no INSA no momento de referência ao estudo dos CI sem tratamento. \*p<0.05 entre casos índice e familiares (T-Test ou teste Mann-Whitney).

#### III.3.2.4 Análise do número de fatores de risco cardiovascular presentes no grupo adultos: indivíduos com FH vs indivíduos sem FH

Também para o grupo adultos, foi realizada uma comparação do número de fatores de risco cardiovascular presentes nos adultos com uma alteração definitiva ou provável vs sem uma alteração, para além da hipercolesterolemia. Verificou-se que existem diferenças estatisticamente significativas entre indivíduos com FH e sem FH, independentemente do número de fatores de risco que apresentem (p=0,002).

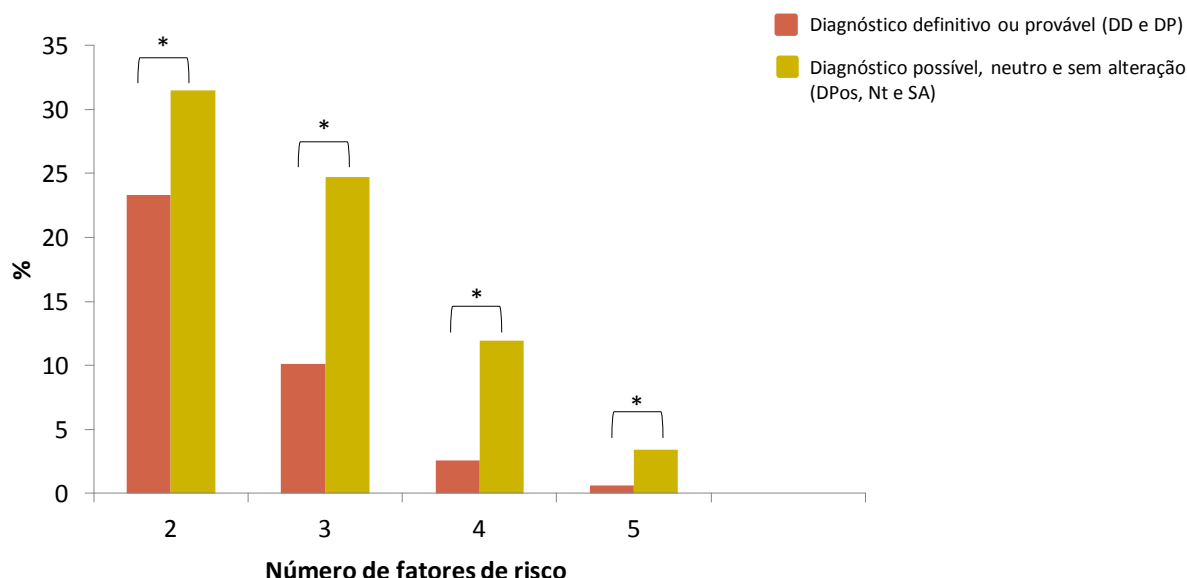


Figura III.3.6 – Número de fatores de risco no grupo adulto. \* Diferença estatisticamente significativa entre os grupos para a presença de fatores de risco (Teste  $\chi^2$ ).

### III.3.3 Caracterização clínica e bioquímica dos doentes com FH homozigótica e heterozigótica composta

O estudo português de hipercolesterolemia familiar identificou 3 doentes homozigóticos e 5 heterozigóticos compostos com mutações no *LDLR* embora alguns indivíduos apresentem alterações cujo estudo funcional não foi realizado. As características clínicas e bioquímicas encontram-se na tabela III.3.9.. De forma a classificar as alterações foi realizada uma análise *in silico* baseada em 6 programas bioinformáticos e o resultado encontra-se na tabela III.3.9. Após a aplicação do critério de classificação das alterações verificou-se que apenas 2 CI são classificados como homozigótico e 1 CI como heterozigótico composto. Os estudos funcionais das alterações apresentadas pelos restantes CI serão importante para a classificação final dos doentes.

Tabela III.3.9 – Características clínicas e bioquímicas dos homozigóticos do EPHF.

Sexo Idade	DCV	Xantomas	Perfil lipídico básico (at/dt) mg/dL			Tratamento		Alterações		Classificação alteração	Classificação final do doente
			CT	cLDL	cHDL	TG		cDNA	Proteína		
F 55	Não	Não	---/299	---/232	---/47	---/99	Estatina	[c.1285G>C] + [c.1285G>C]	[p.Val429Ileu] + [p.Val429Leu]	Patogénica ‡	Homozigótico definitivo
F 36	Não	Não	490/355	435/280-	42/51	64/219	Estatina + ezetimiba	[c.313+6C>T] + [c.1291G>A]	[c.313+6C>T] + [p.Ala431Thr]	Patogénica ‡	Heterozigótico composto definitivo
F 25	Não	Não	596/370	512/299	65/54	96/84-	Estatina	[c.1291G>A] + [c.1291G>A]	[p.Ala431Thr] + [p.Ala431Thr]	Patogénica ‡	Homozigótico definitivo
M 29	EAM (23)	Não	561/346	515/298	---/29	---/96	Estatina + ezetimiba + LDL aferese	([c.631C>G] + [c.1816C>T])* + [c.1178delA])	([p.His211Asp] + [p.Ala606Ser])* + + [p.Lys393ArgfsX20]	Possível ¥ Possível ¥ Patogénica ‡	Heterozigótico composto possível
M 55	EAM (48)	Xantasma	---/409	---/299	---/37	---/363	Estatina + ezetimiba	[c.670G>A]	[p.Asp224Asn] + [p.Glu716Lys]	Patogénica ‡ Possível ¥	Heterozigótico composto possível
F 21	Não	Não	345/258	210/186	51/50	---/79	Estatina + ezetimiba	[c.1216C>T]	[p.Arg406Trp] + [p.Arg406Trp]	Possível ¥	Homozigótico possível
M 46	Não	Não	600/185	---/108	---/63	---/36	Estatina + ezetimiba + Ac Nic	[c.1633G>T]	[p.Gly545Trp] + [p.Gly592Glu]	Possível ¥ Patogénica ‡	Heterozigótico composto possível
F 61	Angina (50)	Não	587/389	401/232	69/62	67/112	Estatina + ezetimiba + LDL aferese	[c.1060-7_1705+7del] + [c.1216C>T]	[Ex8_Ext2del] + [p.Arg406Trp]	Patogénica ‡ Possível ¥	Heterozigótico composto possível

DCV: doença cardiovascular; (at/dt), (antes tratamento/ depois tratamento); CT: colesterol total; cLDL: colesterol de baixa densidade; cHDL: colesterol das lipoproteínas de alta densidade; TG: triglicéidos, EAM: enfarte agudo do miocárdio. \* No mesmo alelo ‡ com estudos funcionais; ¥ classificação baseada em análise *in silico*; - valores desconhecidos





## Parte 4– Sequenciação de exoma





## Parte 4 – Sequenciação de exoma

Durante os últimos anos assistiu-se a um grande desenvolvimento tecnológico no mapeamento do genoma humano, através de técnicas de sequenciação em larga escala. Contudo estas técnicas lançaram também um desafio no modo de análise e gestão dos dados obtidos uma vez que geram bilhões de dados. Foi necessário desenvolver infraestruturas de sistemas de informação, armazenamento de dados, controle de qualidade, análise computacional para leitura e organização de dados, bem como, gestão de informação de dados laboratoriais para validação e controlo das sequências obtidas e gestão de processos, de modo a se conseguir analisar todos os dados (Yang *et al.* 2009; Metzker 2010; Kuhlenbäumer *et al.* 2011; Gilissen *et al.* 2012).

Foram selecionados 5 casos índice do Estudo Português de Hipercolesterolemia Familiar com diagnóstico clínico de FH para o estudo de todo o exoma através de NGS. A seleção destes CI foi efetuada tendo em conta o fenótipo agressivo apresentado pelos CI e familiares hipercolesterolémicos. Um dos critérios utilizados foi a existência de familiares hipercolesterolémicos para todos os CI de forma a se poder realizar o estudo de co-segregação das alterações nas famílias após o estudo da sequenciação de exoma. As características clínicas e bioquímicas dos 5 CI encontram-se na tabela III.4.1. No CI 10146 tinha sido identificada uma mutação no gene *LDLR*, no entanto uma vez que apresentava um fenótipo severo, foi decidido realizar o estudo por NGS. O CI 28120 apresentava uma alteração não descrita anteriormente no gene *APOB*, para a qual não existiam ainda estudos funcionais e a co-segregação da alteração na família parecia não ser total. Os restantes CI (25043, 26130 e 10288) não apresentavam alteração em nenhum dos três genes estudados no EPHF através das técnicas de biologia molecular utilizadas para o diagnóstico de FH.

Numa primeira análise foi verificado para todos os CI a existência de alterações nos genes *LDLR*, *APOB* e *PCSK9*. Verificou-se que o CI 26130 apresentava a alteração p.Ile441Thr no exão 9 do gene *LDLR*. Esta alteração não tinha sido detetada quando foi feito o estudo ao caso índice no âmbito do EPHF. De forma a procurar uma explicação para este caso, foi numa primeira fase, averiguada a sequência referente ao fragmento do exão 9 e 10 realizada no âmbito do EPHF, verificando-se que a alteração não estava presente, nem existia nenhum polimorfismo no local de hibridação de um dos *primers* que poderia estar a impedir o emparelhamento normal. Numa segunda fase e de modo a

esclarecer a ausência da alteração na amostra do CI 26130 foi realizada uma amplificação por PCR do fragmento referente ao exão 9 do gene *LDLR* em duas amostras diferentes do CI 26130, a utilizada para a realização do estudo no âmbito do EPHF e a utilizada para o estudo de *exome sequencing*. Verificou-se que a amostra estudada por *exome sequencing* tinha a alteração enquanto a amostra mais antiga não apresentava a alteração. Possivelmente ocorreu uma troca no momento da colheita ao CI, não pertencendo a primeira amostra ao mesmo. O estudo familiar permitiu detetar a presença da alteração nos familiares hipercolesterolémicos que apresentavam a alteração.

Tabela III.4.1 – Características clínicas e bioquímicas dos CI estudados por *exome sequencing*

Amostra	Idade/ Sexo	CT	cLDL	APOB	Tratamento	DCV	História Familiar	
							CT↑	DCV
10146	11/M	416	334	122	Estatina	Não	Sim	Sim
25043	10/F	308	171	113	Não	Não	Sim	Não
26130	73/F	481	394	150*	Estatina e Ezetimiba	Não	Sim	Sim
28120	21/M	256	191	119	Estatina	Não	Sim	Não
10288	67/M	231*	125*	120*	Estatina e Ezetimiba	Sim	Sim	Sim

M, masculino; F, feminino; CT↑, colesterol total elevado; cLDL, colesterol LDL. Os parâmetros bioquímicos encontram-se em mg/dL; DCV, doença cardiovascular

O estudo de todo o exoma por NGS originou uma elevada quantidade de dados que foram fornecidos pela empresa prestadora do serviço. Para análise foram consideradas todas as alterações que passaram o controlo de qualidade, ou seja, foram apenas consideradas para análise as sequências corretamente amplificadas e com o tamanho esperado. Foram considerados para análise todos os fragmentos que apresentavam uma cobertura superior a sete, ou seja, existiam pelo menos 7 fragmentos legíveis para análise. Por cada amostra obteve-se cerca de 18000 alterações em todo o exoma (tabela III.4.2).

Uma vez que se verificou uma elevada quantidade de dados obtidos para cada amostra, optou-se por analisar os dados com base no pressuposto que as mutações associadas à FH são em geral raras. Assim, optou-se por analisar apenas as alterações que surgiam apenas uma vez em todas as amostras. Os resultados obtidos para cada CI encontram-se na tabela III.4.3.

Tabela III.4.2 – Alterações encontradas em cada caso índice estudado sequenciação de exoma.

Amostra	Número de genes	Número de alterações	Não descritas
10146	8464	18018	855
25043	8440	17958	986
26130	8354	17859	903
28120	8497	18143	938
10288	8462	18176	959

Tabela III.4.3 – Alterações únicas encontradas em cada caso índice estudado por sequenciação de exoma.

Número total de alterações detetadas		Amostra				
		10143	25043	26130	28120	10288
Tipo de alterações	Geral	2408	2546	2280	2491	2432
	Não sinónimas	1168	1219	326	1274	1216
	Sinónimas	1145	1190	209	1114	1123
	<i>Splicing</i>	26	31	8	36	30
	<i>Frameshift</i>	30	37	30	28	21
	<i>Indel</i>	24	49	46	21	32
	<i>Stop</i>	15	20	11	18	10
	Não descritas	516	630	565	568	613
Tipo de alterações	Não sinónimas	283	275	310	303	340
	Sinónimas	169	249	188	201	211
	<i>Splicing</i>	5	36	5	8	8
	<i>Frameshift</i>	27	36	34	27	19
	<i>Indel</i>	22	19	18	18	29
	<i>Stop</i>	10	15	10	11	6

Uma vez que o número de alterações era ainda elevado (mais de 500 alterações por amostra) e com base nos pressupostos descritos anteriormente, foi elaborada uma

estratégia que reduzisse o número de alterações a analisar por sequenciação de Sanger e levasse até à identificação de uma possível alteração que fosse responsável pela hipercolesterolemia na família. Assim a abordagem escolhida foi selecionar genes descritos anteriormente associados a níveis elevados de colesterol total e/ou colesterol LDL, com base em resultados descritos na literatura, principalmente em diversos estudos de *genome whole association studies* (GWAS) ou estudos de *linkage* (Coon *et al.* 2001; Coon *et al.* 2002; Sandhu *et al.* 2008; Willer *et al.* 2008). Foram selecionados 42 genes, incluindo *LDLR*, *APOB*, *PCSK9* e *LDLRAP1* (tabela III.4.4). Outra abordagem complementar, foi verificar no programa Cytoscape, um programa de interação de proteínas, quais as proteínas que já tinham sido descritas como estando a interagir com a *LDLR* (Figura III.4.1 e tabela III.4.4), tendo-se identificado e selecionado para estudo mais 7 genes cujas suas proteínas interagem com *LDLR*, uma vez que as restantes proteínas já tinham sido incluídas na seleção de genes associados a estudos de *linkage* e GWAS.

Após a seleção dos genes foi verificado nas alterações que surgiam apenas uma vez em todas as amostras se algum dos 49 genes apresentava uma alteração nas amostras estudadas por *exome sequencing*, tendo-se encontrado 50 alterações em 22 dos 49 genes selecionados (Figura III.4.2). Das 50 alterações, 30 eram sinónimas e o estudo de co-segregação nas famílias não foi realizado no decorrer deste estudo. As restantes 20 alterações encontram-se distribuídas por 17 genes, incluindo *LDLR*, *APOB* e *LDLRAP1*, tendo sido verificadas nos CI e seus familiares (Tabela III.4.5).

Tabela III.4.4 – Genes selecionados para pesquisa de alterações descritas em estudos prévios de associação (GWAS), *linkage* e no programa CYTOSCAPE.

Nome do gene	Nome da proteína	Locus	Tipo de estudo	Referência
<b>DHCR7</b>	7-dehydrocholesterol reductase	11q13	Estudos de <i>linkage</i>	Coon H <i>et al</i> , 2002
<b>OSBP</b>	Oxysterol-binding protein 1	11q12	Estudos de <i>linkage</i>	Coon H <i>et al</i> , 2002
<b>ABCB10</b>	ATP-binding cassette sub-family B member 10, mitochondrial	1q42	Estudos de <i>linkage</i>	Coon H <i>et al</i> , 2002
<b>GGPS1</b>	Geranylgeranyl pyrophosphate synthase	1q43	Estudos de <i>linkage</i>	Bossé <i>et al</i> , 2004
<b>LRP5</b>	Low-density lipoprotein receptor-related protein 5	11q13	Estudos de <i>linkage</i>	Bossé <i>et al</i> , 2004
<b>CPT1A</b>	Carnitine O-palmitoyltransferase 1, liver isoform	11q13	Estudos de <i>linkage</i>	Bossé <i>et al</i> , 2004
<b>UCP2</b>	Mitochondrial uncoupling protein 2	11q13	Estudos de <i>linkage</i>	Bossé <i>et al</i> , 2004
<b>UCP3</b>	Mitochondrial uncoupling protein 3	11q13	Estudos de <i>linkage</i>	Bossé <i>et al</i> , 2004
<b>SOAT1/ACAT1</b>	Sterol O-acyltransferase 1	1q25	Estudos de <i>linkage</i>	Bossé <i>et al</i> , 2004
<b>APOA1</b>	Apolipoprotein A-I	11q23	Estudos de <i>linkage</i>	Bossé <i>et al</i> , 2004
<b>APOC3</b>	Apolipoprotein C-III	11q23	Estudos de <i>linkage</i>	Bossé <i>et al</i> , 2004
<b>APOA4</b>	Apolipoprotein A-IV	11q23	Estudos de <i>linkage</i>	Bossé <i>et al</i> , 2004
<b>APOA5</b>	Apolipoprotein A-V	11q23	Estudos de <i>linkage</i>	Bossé <i>et al</i> , 2004
<b>CYP11A</b>	Cholesterol side-chain cleavage enzyme, mitochondrial	15q23-q24	Estudos de <i>linkage</i>	Bossé <i>et al</i> , 2004
<b>LRP3</b>	Low-density lipoprotein receptor-related protein 3	19q13	Estudos de <i>linkage</i>	Bossé <i>et al</i> , 2004
<b>LIPE</b>	hormone-sensitive lipase	19q13	Estudos de <i>linkage</i>	Bossé <i>et al</i> , 2004
<b>APOC1</b>	Apolipoprotein C-I	19q13	Estudos de <i>linkage</i>	Bossé <i>et al</i> , 2004
<b>APOC2</b>	Apolipoprotein C-II	19q13	Estudos de <i>linkage</i>	Bossé <i>et al</i> , 2004
<b>APOC4</b>	Apolipoprotein C-IV	19q13	Estudos de <i>linkage</i>	Bossé <i>et al</i> , 2004
<b>APOE</b>	Apolipoprotein E	19q13	Estudos de <i>linkage</i>	Bossé <i>et al</i> , 2004
<b>PSRC1</b>	Proline/serine-rich coiled-coil protein 1	1p13	GWAS	Willer <i>et al</i> , 2008; Sandhu <i>et al</i> 2008
<b>CELSR2</b>	Cadherin EGF LAG seven-pass G-type receptor 2	1p13	GWAS	Willer <i>et al</i> , 2008; Sandhu <i>et al</i> 2008
<b>SORT1</b>	Sortilin	1p13	GWAS	Willer <i>et al</i> , 2008
<b>NCAN</b>	Neurocan core protein	19p12	GWAS	Willer <i>et al</i> , 2008
<b>CHLP2</b>	Cartilage intermediate layer protein 2	19p13	GWAS	Willer <i>et al</i> , 2008
<b>B3GALT4</b>	Beta-1,3-galactosyltransferase 4	6p21	GWAS	Willer <i>et al</i> , 2008
<b>B4GALT4</b>	Beta-1,4-galactosyltransferase 4	3q13	GWAS	Willer <i>et al</i> , 2008

Tabela III.4.4 – continuação

Nome do gene	Nome da proteína	Locus	Tipo de estudo	Referência
<b>ABCG8</b>	ATP-binding cassette sub-family G member 8	2p21	GWAS	Domitrescu <i>et al</i> 2011
<b>PBX4</b>	Pre-B-cell leukemia transcription factor 4	19p13	GWAS	Kathiresan <i>et al</i> , 2008
<b>HMGCR</b>	3-hydroxy-3-methylglutaryl-coenzyme A reductase	6q13	GWAS	Kathiresan <i>et al</i> , 2008
<b>TIMD4</b>	T-cell immunoglobulin and mucin domain-containing protein 4	5q33	GWAS	Domitrescu <i>et al</i> 2011
<b>HNF1A</b>	Hepatocyte nuclear factor 1-alpha	12q24	GWAS	Domitrescu <i>et al</i> 2011
<b>BCL3</b>	B-cell lymphoma 3 protein	19q13	GWAS	Domitrescu <i>et al</i> 2011
<b>TOMM40</b>	Mitochondrial import receptor subunit TOM40 homolog	19q13	GWAS	Domitrescu <i>et al</i> 2011
<b>ABCG1</b>	ATP-binding cassette sub-family G member 1	21q22	Estudos de <i>linkage</i>	Wang <i>et al</i> 2011
<b>ABCA1</b>	ATP-binding cassette sub-family A (ABC1), member 1	9q31.1	Estudos de <i>linkage</i>	Wang <i>et al</i> 2011
<b>LSS</b>	Lanosterol synthase	21q22	Estudos de <i>linkage</i>	Wang <i>et al</i> 2011
<b>GCKR</b>	Glucokinase regulatory protein	2p23	GWAS	Chambers <i>et al</i> , 2011
<b>MTTP</b>	Microsomal triglyceride transfer protein large subunit	4q24	PPI networks	
<b>SMAD3</b>	Mothers against decapentaplegic homolog 3	15q22.33	PPI networks	
<b>LRP2</b>	low density lipoprotein receptor-related protein 2	2q24-q31	PPI networks	
<b>LRPAP1</b>	Low Density Lipoprotein Receptor-Related Protein Associated Protein 1	4p16.3	PPI networks	
<b>PDIA3</b>	protein disulfide isomerase family A, member 3	15q15	PPI networks	
<b>DAB2</b>	disabled homolog 2, mitogen-responsive phosphoprotein	5p13	PPI networks	
<b>FLT1</b>	Vascular endothelial growth factor receptor 1 precursor	13q12.3	PPI networks	

GWAS, Genome Wide Association Studies; PPI networks, protein-protein interaction (ppi) networks.

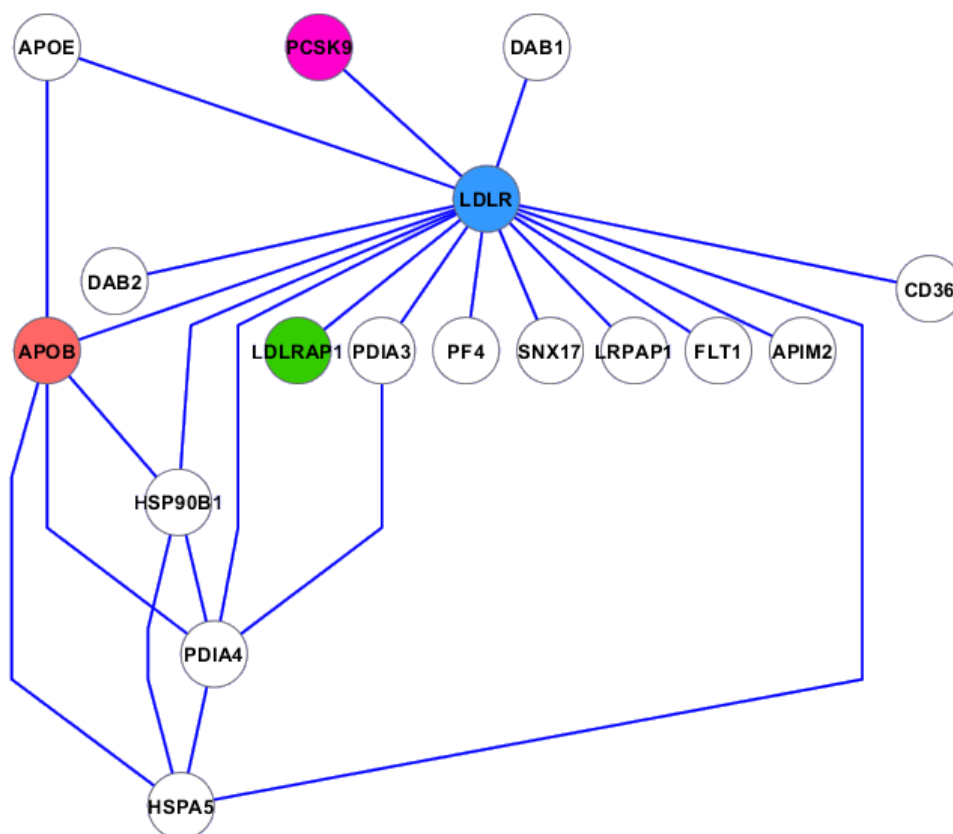


Figura III.4.1 – Interação de proteínas com a proteína LDLR obtida através do programa Cytoscape. As proteínas com cores são as descritas anteriormente como causadores de FH.

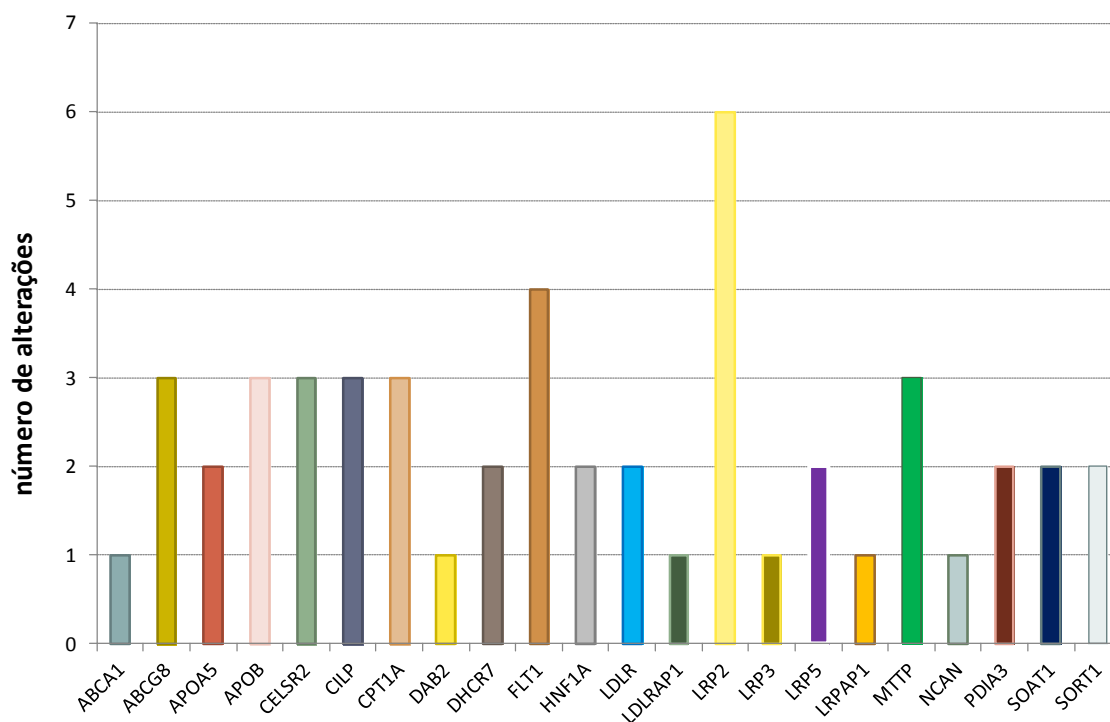


Figura III.4.2 – Distribuição das 50 alterações encontradas em 22 dos genes selecionados por estudos de *linkage*, GWAS e através do programa Cytoscape.

#### IV.1 Caso índice 10146

O CI 10146 foi incluído no estudo de exoma apesar de apresentar a mutação p.Ala431Thr no gene *LDLR*, contudo, uma vez que apresentava um fenótipo agressivo e ambos os pais apresentavam critérios clínicos de FH, dado que a mãe apresentava unicamente a mutação p.Ala431Thr no gene *LDLR* e o pai não apresentava qualquer alteração num dos genes associados à FH. O pai do CI foi inclusivamente estudado por pirosequenciação também, não tendo sido detetada qualquer alteração. Desta forma, optou-se pelo estudo completo do exoma de forma a melhor clarificar o seu fenótipo. O CI apresentava a mutação no gene *LDLR* herdada da mãe, uma vez que o pai apresentava critérios clínicos de FH podendo ter uma alteração noutra gene que justificasse hipercolesterolemia. Pela análise dos resultados do *exome sequencing* verificou-se que o CI 10146 apresentava alterações nos genes *SOAT1*, *CELSR2*, *FLT1*, *LRPAP1*, *PDIA3*, *ABCG8* e *LRP2* (tabela III.4.5). De forma a verificar a co-segregação das alterações com a hipercolesterolemia na família foram realizados estudos moleculares a todos os indivíduos da família tendo-se verificado que, das 9 alterações estudadas, o pai (II:1) do CI apresentava apenas a alteração p.Gln526Arg no gene *SOAT1* em homozigotia dado que as restantes 8 alterações estavam presentes na mãe do CI (II:2) (Figura III.4.3). A alteração p.Gln526Arg no gene *SOAT1* parece co-segregar na família com a hipercolesterolemia uma vez que todos os indivíduos hipercolesterolémicos paternos apresentam a alteração. A alteração p.Arg1045Gln no gene *FLT1* encontra-se presente na mãe, avô materno e irmão do CI e co-segregando assim na família materna com a hipercolesterolemia. A mutação p.Ala431Thr no gene *LDLR* está presente nos mesmos indivíduos, no entanto todos apresentavam um fenótipo agressivo, podendo a alteração no gene *FLT1* ser a responsável por este fenótipo. A análise da alteração através de 6 programas bioinformáticos previu que a alteração seria patogénica, no entanto apenas os estudos funcionais poderão comprovar a sua patogenicidade. Durante o decorrer deste trabalho não foi possível realizar os estudos funcionais das alterações encontradas nesta família. As alterações nos genes *CELSR2*, *LRPAP1*, *PDIA3*, *ABCG8* e *LRP2* foram detetadas no CI, irmão, mãe e avó materna. Estas alterações parecem não co-segregar com a hipercolesterolemia na família, uma vez que a avó materna apresentava uma boa resposta a terapêutica farmacológica instituída, apresentando níveis de colesterol total e LDL dentro das recomendações da Sociedade Europeia de Cardiologia (Catapano *et al.* 2011b), o que não é usual em doentes com FH. O avô materno encontrava-se também a cumprir



terapêutica farmacológica hipolipemiante apresentado ainda valores elevados como é característica de doentes com FH com uma mutação no gene *LDLR* que diminuem uma atividade no recetor das LDL (Figura III.4.3).

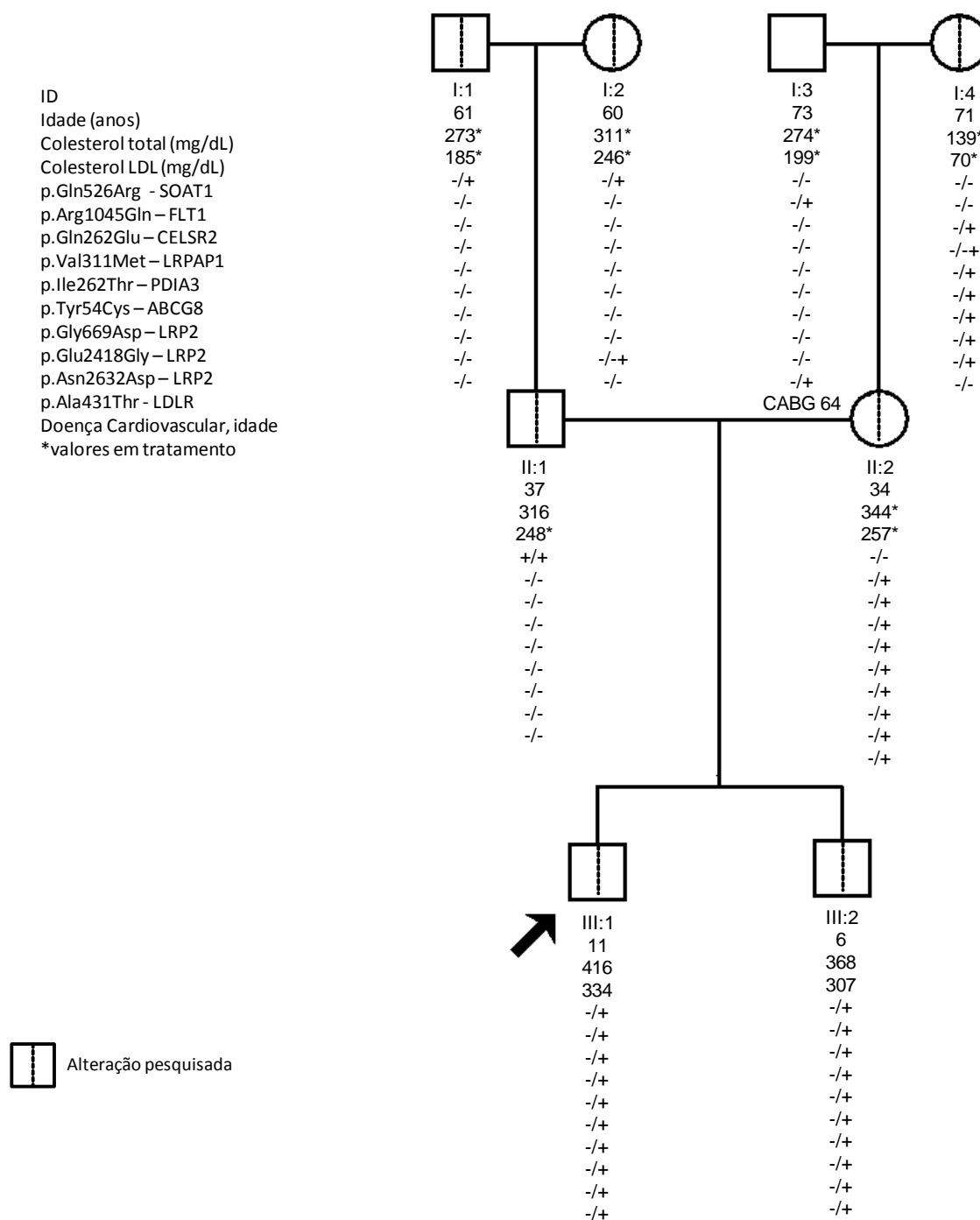


Figura III.4.3 – Árvore genealógica do CI 10146 (III.1). Pesquisa de alterações nos genes *SOAT1*, *FLT1*, *CELSR2*, *LRPAP1*, *PDIA3*, *ABCG8* e *LRP2* detetadas por *exome sequencing*. A seta representa o caso índice. CABG, Coronary Artery Bypass Grafting (Bypass da artéria coronária).

## IV.2 Caso índice 25043

Após o estudo do exoma verificou-se que o caso índice 25043 apresentava 4 alterações nos genes *CPT1A*, *FLT1*, *OSBPL1A*, *SMAD3* e *SORT1*. Foi realizado o estudo molecular a todos os indivíduos da família que as alterações p.Arg224Gln no gene *FLT1* e p.Glu444Gln no gene *SORT1* encontram-se presente no CI, no pai e na tia do CI, no entanto não está presente nas primas (Figura III.4.4). Não são conhecidos os valores de CT e cLDL do pai desta familiar, podendo a hipercolesterolemia neste caso ter sido herdada do pai. A alteração p.Cys39del no gene *OSBPL1A* encontra-se presente na mãe normolipidémica e no pai hiperlipidémico do CI, não devendo por este motivo ser a causa de doença. A alteração. De acordo com os programas bioinformáticos para previsão *in silico*, as alterações nos genes *SORT1* e *FLT1* foram consideradas possivelmente patogénicas (tabela III.4.5), sendo em ambos os casos importantes realizar o estudo funcional das alterações de forma a comprovar ou não a sua patogénicidade.

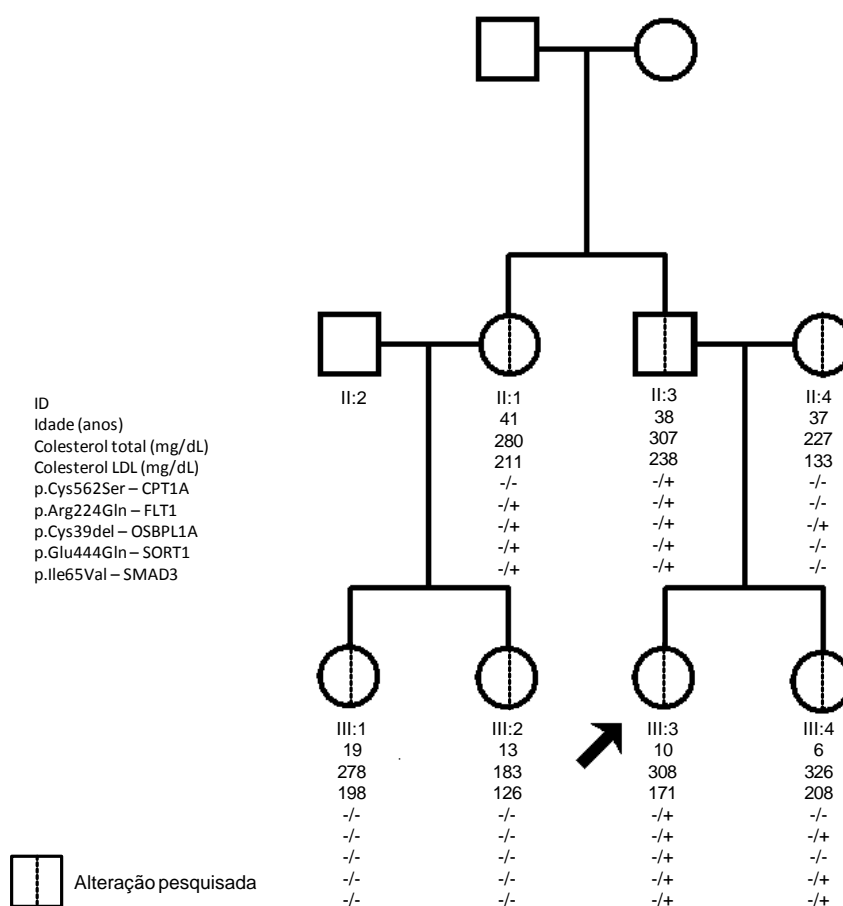


Figura III.4.4 – Árvore genealógica do caso índice 25043 (III.3). Pesquisa de alterações nos genes *CPT1A*, *FLT1* e *OSBPL1A* detetadas por *exome sequencing*. A seta representa o caso índice. Os símbolos em branco são referentes a familiares não analisados.

### IV.3 Caso índice 28120

O estudo do exoma ao CI 28120 detetou a existência de alterações nos genes *LRP5*, *ABCG8* e *DAB2* (tabela III.4.5), para além da mutação no gene *APOB* detetada anteriormente por pirosequenciação. Optou-se por realizar o estudo do exoma a este CI uma vez que a alteração no gene *APOB* está localizada fora da região de ligação ao recetor das LDL e não co-segregava totalmente na família com a hipercolesterolemia. Quando foi realizado o estudo de sequenciação do exoma o estudo funcional não tinha sido ainda efetuado. Como referido anterior, foi encontrada a alteração p.Asp19His no gene *ABCG8* tendo-se verificado após análise à história clínica dos familiares que a avó materna do CI fez um colecistectomia (devido à patologia biliar) justificando a relação fenótipo/genótipo na avó. Uma vez que os restantes indivíduos onde a alteração foi também detetada são mais jovens os sintomas relacionados com as litíases biliares podem ainda não ter surgido. Neste CI foram também encontradas as alterações p.Val667Met e p.Arg296Cys nos genes *LRP5* e *DAB2*, respetivamente. O estudo molecular aos familiares do CI 28120 verificou que ambas as alterações encontravam-se presente na mãe, avô materno, tio e primas (Figura III.4.5). Encontra-se descrito que a proteína Dab2 pode ser uma das proteínas responsáveis em algumas células, nomeadamente nos fibroblastos, pela formação de vesículas de clatrina, importantes para a internalização do complexo LDLR:LDL (Soutar 2010). Todas as alterações foram analisadas por 6 programas bioinformáticos diferentes de forma a realizar a previsão que cada alteração originava a proteína. Com base no critério estabelecido na parte 2 deste capítulo a alteração p.Arg296Cys no gene *DAB2* é neutra e a alteração no gene *LRP5* provavelmente patogénica. Não foi possível a realização de estudos funcionais a ambas as alterações no âmbito deste trabalho.

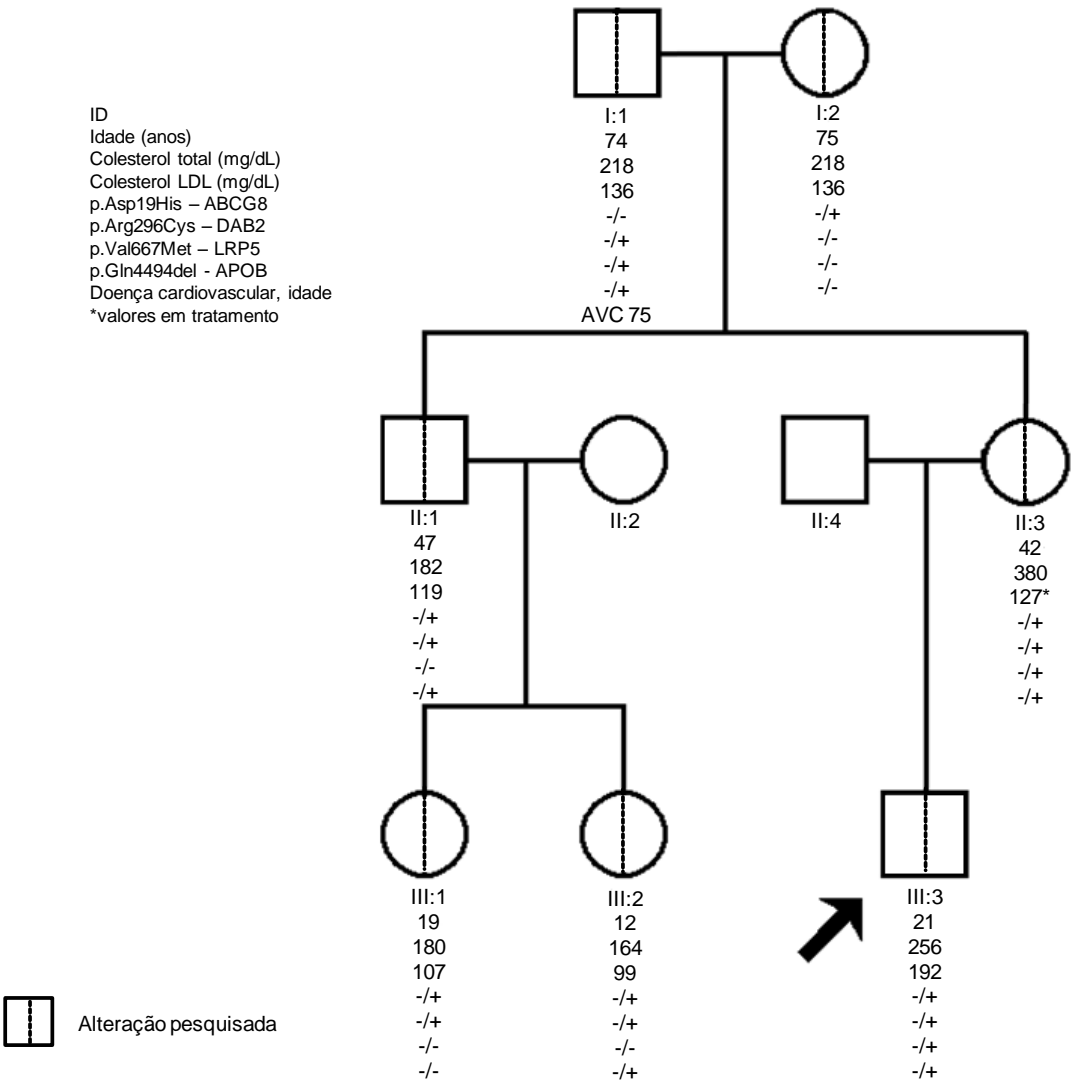


Figura III.4.5 - Árvore genealógica do caso índice 28120 (III.3). Estudo familiar para a pesquisa de alterações nos genes *ABCG8*, *DAB2* e *LRP5* detetadas por *exome sequencing*. A seta representa o caso índice. AVC, acidente vascular cerebral. Os símbolos em branco são referentes a familiares não analisados.

#### IV.4 Caso índice 10288

O caso índice 10288 apresentava um fenótipo típico de FH, com doença cardiovascular. No entanto após o estudo aos genes *LDLR*, *PCSK9* e *APOB* não foi detetada nenhuma alteração. O estudo do exoma neste CI identificou 3 alterações nos genes *CELSR2*, *LDLRAP1* e *MTTP*. As alterações p.Arg238Trp e p.Gln244Glu nos genes *LDLRAP1* e *MTTP*, foram encontradas apenas no CI. Após o estudo molecular da alteração p.Ala2689Ser no gene *CELSR2* verificou-se não existir co-segregação com a hipercolesterolemia na família (Figura III.4.6). A análise das alterações p.Ala2689Ser e

p.Gln244Glu nos genes *CELSR2* e *MTTP*, respetivamente, através de programas bioinformáticos previu que as mesmas não são patogénicas (tabela III.4.5). A alteração p.Arg238Trp no gene *LDLRAP1* encontra-se descrita na literatura como sendo um polimorfismo (Soutar and Naoumova 2004), não sendo claro que não influencie os níveis plasmáticos de cLDL.

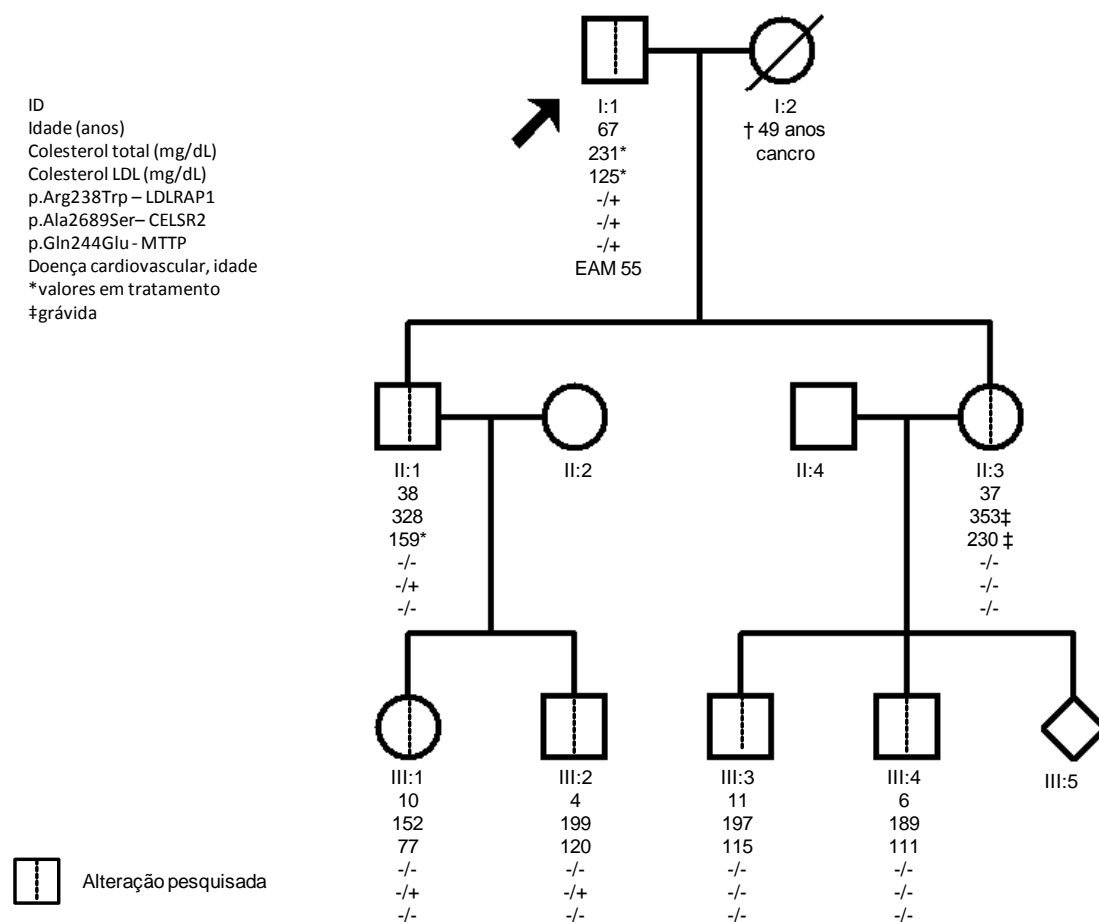


Figura III.4.6 - Árvore genealógica do caso índice 10288 (III.3). Pesquisa de alterações nos genes *LDLRAP1*, *CELSR2* e *MTTP* detetadas por *exome sequencing*. A seta representa o caso índice. EAM, enfarte do miocárdio. Os símbolos em branco são referentes a familiares não analisados.

Tabela III.4.5 – Caracterização das alterações encontradas nos 4 doentes do EPHF onde foi realizado a sequenciação do

Gene	Alteração		Frequência				Análise <i>in silico</i>				Previsão		Amostra
	cDNA	Proteína	1000G	ESP	SIFT	PolyPhen2	Condel	Mutation Taster	Grantham Score	PhyloP	<i>in silico</i>		
LRP5	c.1999G>A	p.Val667Met	0,02%	0,04%	Patogénica	Patogénica	Neutra	Causa doença	Conservado	4,58	Possível	rs4988321	28120*
CPT1A	c.1685G>C	p.Cys562Ser	nd	nd	Neutra	Neutra	Neutra	Causa doença	Moderado conservado	2,14	Neutra	nd	25043
SOAT1	c.1577T>G	p.Gln526Arg	0,12%	0,05%	Neutra	Neutra	Neutra	Causa doença	Conservado	2,89	Neutra	rs13306731	10146
CELSR2	c.784C>G	p.Gln262Glu	nd	nd	Patogénica	Neutra	Neutra	Polimorfismo	Conservado	1,66	Neutra	nd	10146
	c.8065G>T	p.Ala2689Ser	nd	nd	Neutra	Neutra	Neutra	Polimorfismo	Moderado conservado	0,85	Neutra	nd	10288
SORT1	c.919G>C	p.Glu444Gln	0,01%	0,01%	Patogénico	Neutra	Neutra	Causa doença	Conservado	5,49	Possível	rs2228606	25043
ABCG8	c.55G>C	p.Asp19His	0,07%	0,05%	Patogénico	Neutra	Neutra	Polimorfismo	Moderado	1,47	Neutra	rs11887534	28120* <sup>1</sup>
	c.161A>G	p.Tyr54Cys	44,10%	32%	Patogénico	Neutra	Neutra	Polimorfismo	Radical	2,95	Neutra	rs4148211	10146
FLT1	c.671G>A	p.Arg224Gln	nd	nd	Patogénico	Patogénico	Neutra	Causa doença	Moderado	4	Possível	nd	25043
	c.3134G>A	p.Arg1045Gln	nd	nd	Patogénico	Patogénico	Patogénico	Causa doença	Conservado	6,26	Provável	rs139361304	10146
LRPAP1	c.931G>A	p.Val311Met	0,07%	0,01%	Neutra	Possível	Neutra	Polimorfismo	Conservada	0,41	Neutra	rs1800493	10146
PDIA3	c.785T>C	p.Ile262Thr	nd	nd	Neutra	Neutra	Neutra	Polimorfismo	Moderado conservado	0,95	Neutra	nd	10146
DAB2	c.886C>T	p.Pro296Cys	0,08%	0,03%	Neutra	Neutra	Neutra	Causa doença	Radical	1,98	Neutra	rs79344176	28120
OSBPL1A	c.115_116insAATT	p.Cys39fs*1	nd	0,27%	nd	nd	nd	Causa doença	nd	3,8	Provável	rs74793804	25043
MTTP	c.730C>G	p.Gln244Glu	0,02%	0,04%	Neutra	Neutra	Neutra	Polimorfismo	Conservado	0,29	Neutra	rs17599091	10288
SMAD3	c.193A>G	p.Ile65Val	nd	nd	Neutra	Neutra	Neutra	Causa doença	Conservado	1,17	Neutra	nd	25043
LRP2	c.2006G>A	p.Gly669Val	0,02%	0,03%	Patogénico	Patogénico	Patogénico	Causa doença	Moderado conservado	6,1	Provável	rs34291900	10146
	c.7253A>G	p.Glu2418Gly	0,07%	0,01%	Neutra	Neutra	Neutra	Polimorfismo	Moderado	0,69	Neutra	rs61995915	10146
	c.7894A>G	p.Asn2632Asp	0,02%	0,03%	Neutra	Patogénico	Neutra	Polimorfismo	Conservado	2,47	Neutra	rs17848169	10146
LDLRAP1	c.712C>T	p.Arg238Trp	2,40%	3,60%	Patogénico	Patogénico	Patogénico	Polimorfismo	Moderado conservado	-0,04	Possível	rs41291058	10288

ESP, exome sequencing project; 1000G, 1000 Genome Project; dbSNP – base de dados SNP; nd – informação não disponível; \* alteração descrita na OMIM como originando litíase biliar; \*<sup>1</sup> alteração descrita na OMIM como originando síndrome *osteoporosis-pseudoglioma*.







## CAPÍTULO IV

---

### DISCUSSÃO



## IV DISCUSSÃO

A Hipercolesterolemia Familiar é uma das doenças monogénicas mais bem estudada no mundo e apresenta critérios clínicos bem definidos (Marks *et al.* 2003). Mutações nos genes *LDLR*, *APOB* e *PCSK9* estão associadas a uma forma dominante da doença e mutações no gene *LDLRAP1* à forma recessiva da doença (Soutar and Naoumova 2007). No entanto, estudos realizados em diferentes *cohorts* indicam que 20-80% dos doentes com diagnóstico clínico de FH não apresentam mutações em nenhum destes genes (Fouchier *et al.* 2001; Ikonen 2006; Leren *et al.* 2008; Nordestgaard *et al.* 2013; Talmud *et al.* 2013), estando esta diferença associada aos diferentes critérios de diagnóstico e seus *cut off* aplicados nos diferentes grupos de trabalho. No estudo português de Hipercolesterolemia Familiar (EPHF), até ao ano de 2010, apenas 48% dos doentes com diagnóstico clínico de FH apresentavam uma mutação nos genes *LDLR*, *APOB* e *PCSK9* (Medeiros *et al.* 2010), pelo que a causa genética nos restantes doentes pode estar em mutações nestes genes mas situadas em regiões habitualmente não estudadas ou em alterações noutros genes do metabolismo lipídico. Além disso, é também possível que alguns destes doentes tenham uma dislipidemia poligénica e não monogénica, como referido recentemente por Talmud e colegas (Talmud *et al.* 2013).

### IV.1 Caracterização molecular de casos índice com diagnóstico clínico de FH no âmbito do EPHF

Ao longo de 13 anos, o EPHF identificou molecularmente 294 casos índice com uma alteração patogénica ou potencialmente patogénica num dos genes associados à FH. Através dos estudos familiares, foi possível identificar-se mais 339 doentes com FH.

Em 2008, data do início deste trabalho, o diagnóstico molecular da FH era realizado por cromatografia líquida de alta pressão desnaturante (DHPLC), PCR e sequenciação de Sanger. Noutros laboratórios, no mundo, o estudo molecular da FH era realizado nesta altura por outros métodos de diagnóstico clássicos como polimorfismo de conformação de cadeia única (SSCP), eletroforese em gel de gradiente desnaturante (DGGE), sequenciação direta, sendo além disso a identificação de grandes rearranjos efetuada apenas nalguns laboratórios por *Southern Blot*. No entanto, em 2007 apareceram os primeiros trabalhos recorrendo à técnica de MLPA para pesquisa de grandes rearranjos, tendo-se começado a estudar, em todo o mundo, os grandes rearranjos por esta técnica.

Também nesta altura o diagnóstico do *PCSK9* foi incluído no diagnóstico molecular por todos os grupos que estudavam a FH. Sendo assim, em 2008 devido aos novos conhecimentos no ciclo do recetor das LDL e ao aparecimento de novas tecnologias, houve necessidade de reestruturar o diagnóstico molecular no âmbito do EPHF para 4 fases tendo sido esta reestruturação realizado no âmbito deste trabalho.

#### **IV.1.1 Pesquisa de mutações nos genes *LDLR* e *APOB* no âmbito do EPHF**

No final da década de 80 e início da década de 90, surgiram os primeiros estudos moleculares no gene *LDLR* em diversos laboratórios, sendo a percentagem de grandes rearranjos reportados nessa altura similar à percentagem de mutações pontuais (Soutar and Naoumova 2007), uma vez que a técnica de *Southern blotting* era uma das mais utilizadas para a deteção de grandes rearranjos no gene *LDLR*. Com a introdução de novas técnicas de diagnóstico molecular a percentagem de grandes rearranjos causadores de FH diminuiu em relação ao número de alterações pontuais e de *splicing* reportadas, encontrando-se descrito na literatura que grandes rearranjos correspondem a 7 -13% das alterações encontradas (Soutar *et al.* 2003; Tosi *et al.* 2007; Taylor *et al.* 2009; Chiou and Charng 2010). A técnica de MLPA foi otimizada no decorrer deste trabalho, existindo neste momento um procedimento com sondas específicas para detetar grandes rearranjos no gene *LDLR*. Esta técnica é utilizada na maioria dos laboratórios como método de diagnóstico uma vez que se verificou ser o melhor método para a deteção de grandes rearranjos (Wang *et al.* 2005; Tosi *et al.* 2007; Taylor *et al.* 2009; Goldmann *et al.* 2010). Esta abordagem permitiu identificar grandes rearranjos em 6,5% dos CI identificados molecularmente no EPHF, sendo esta percentagem ligeiramente inferior à descrita noutras populações (7-13%) (Soutar *et al.* 2003; Mozas *et al.* 2004; Tosi *et al.* 2007; Taylor *et al.* 2009; Chiou and Charng 2010; Marduel *et al.* 2010; Benn *et al.* 2012). A deleção simultânea das zonas desde o promotor até ao exão 2 e desde o exão 8 até ao 12 encontra-se apenas descrita na população portuguesa (Bourbon *et al.* 2008), tendo a deleção das zonas desde o promotor até ao exão 2 sido reportada noutra população (Garuti *et al.* 1996) mas não associada à deleção do exão 8 até ao exão 12. Este grande rearranjo é uma das mutações mais comuns na população portuguesa, correspondendo a uma percentagem de 6% dos CI com uma mutação causadora de FH identificados no EPHF. O fenótipo dos CI com um grande rearranjo é usualmente mais severo que um heterozigótico com mutações pontuais ou de *splicing*,

observando-se também uma menor redução dos níveis de CT e cLDL quando realizando terapêutica hipolipemiante (Marks *et al.* 2003). A implementação da técnica de MLPA permitiu assim aumentar o número de doentes geneticamente identificados no EPHF.

#### IV.1.2 Pesquisa de mutações no gene PCSK9 no âmbito do EPHF

No âmbito da terceira fase do diagnóstico molecular da FH no EPHF foi implementada a deteção de mutações no gene *PCSK9*. Alterações neste gene estão associadas a dois tipos de fenótipos: hipercolesterolemia e hipocolesterolemia (Cohen *et al.* 2005). Neste trabalho foram estudados 272 casos índice para o gene *PCSK9* - em 82 destes foi efetuado o estudo completo do gene e nos restantes foi realizada a pesquisa de alterações nos 4 exões do *PCSK9* onde previamente haviam sido descritas mutações funcionais associadas à hipercolesterolemia familiar. Esta foi a opção escolhida, uma vez que a amplificação de alguns fragmentos do gene *PCSK9* se revelou muito morosa e exaustiva por se tratar de locais ricos em GC. Por outro lado, o número de mutações descritas no gene *PCSK9* é baixo, desaconselhando o investimento no estudo completo do gene. O estudo completo do gene foi realizado, apenas, nos casos índice com um fenótipo mais agressivo e sem mutação nos genes *LDLR* e *APOB*, de modo a encontrar nestes CI a causa molecular para o fenótipo observado.

A primeira mutação encontrada no gene *PCSK9* em Portugal foi a alteração p.Asp374His em 2 casos índice não relacionados (Bourbon *et al.* 2008). Uma mutação no mesmo codão mas originando um aminoácido diferente (p.Asp374Tyr) foi descrita noutras populações, tendo os estudos funcionais comprovado a sua patogenicidade (Benjannet *et al.* 2004). Em 2009 estudos funcionais comprovaram também a patogenicidade da alteração p.Asp374His (Fasano *et al.* 2009), descrita apenas na população portuguesa. Para além desta mutação, foram identificadas em 3 CI duas alterações potencialmente patogénicas, não descritas anteriormente, p.Ala62Asp e p.Pro467Ala (a alteração p.Ala62Asp foi encontrada em 2 CI relacionados), e uma já descrita, p.Arg237Trp com estudo funcional. Na impossibilidade da realização dos estudos funcionais e de forma a melhor caracterizar e classificar as alterações foram realizadas diferentes análises (previsão *in silico*; estudos de co-segregação na família; e verificação num painel de indivíduos normolipidémicos da população portuguesa quanto à presença ou ausência das alterações). A análise *in silico* classificou as alterações p.Ala62Asp e p.Arg237Trp como potencialmente patogénicas e a alteração

p.Pro467Ala uma variante rara. A análise *in silico* sugere que a alteração p.Pro467Ala não tem consequências funcionais - contudo é importante referir que a análise *in silico* de genes complexos como o *PCSK9* é falível, uma vez que a obtenção da cristalização da proteína é difícil (Cameron *et al.* 2006; Poirier and Mayer 2013). A alteração p.Ala62Asp localiza-se no péptido de sinal e até ao momento não foram descritas mutações associadas ao fenótipo de hipercolesterolemia neste local. Por sua vez a alteração p.Pro467Ala localiza-se no domínio C-terminal do gene *PCSK9*, tendo sido já associadas mutações e variantes que originam um fenótipo hipercolesterolemia nesta região (Kotowski *et al.* 2006; Abifadel *et al.* 2009). A alteração p.Arg237Trp encontra-se já descrita em 3 famílias de origem geográfica diferente nas quais origina fenótipos diferentes (Benjannet *et al.* 2004; Berge *et al.* 2006; Cameron *et al.* 2006; Kotowski *et al.* 2006; Homer *et al.* 2008), contudo o estudo funcional verificou um aumento dos níveis de *LDLR* à superfície de células HepG2 (Cameron *et al.* 2006) e consequentemente uma diminuição dos níveis de cLDL, tendo os autores concluído que a alteração não será patogénica.

Os casos índice identificados no EPHF que apresentavam a mutação p.Asp374His apresentavam uma hipercolesterolemia severa e DCV prematura, comparável ao fenótipo verificado para os HoFH identificados. Estes dados levantaram a hipótese que mutações no gene *PCSK9* apresentavam um fenótipo mais agressivo que mutações pontuais no *LDLR*. Contudo, recentemente na literatura, mutações *missense* associadas a ganho de função no gene *PCSK9* têm sido associadas a fenótipos típicos de HtFH (Abifadel *et al.* 2009) e portanto, menos agressivo do que o verificado nos 2 CI identificados em Portugal (Bourbon *et al.* 2008) e semelhantes ao verificado nos CI que apresentam as alterações p.Ala62Asp e p.Pro467Ala.

As mutações descritas no gene *PCSK9* associadas a um ganho de função e consequentemente um fenótipo de hipercolesterolemia (Cohen *et al.* 2005) são usualmente mutações *missense*. Para além disso, todas as mutações nulas descritas neste gene originam fenótipo de hipocolesterolemia (Abifadel *et al.* 2009) e são mais comuns nos indivíduos de ascendência africana (Cohen *et al.* 2005). Encontram-se também descritos alguns polimorfismos funcionais associados a uma diminuição dos níveis de cLDL em várias populações (Abifadel *et al.* 2009). Na população em estudo neste trabalho, não foram encontradas mutações que causem um fenótipo de hipocolesterolemia. No entanto, foram detetados polimorfismos descritos anteriormente

como originando uma diminuição dos valores de cLDL, p.21\_22InsLeu e p.Arg46Leu, com uma frequência de 36% e 1%, respetivamente. Estes polimorfismos foram também encontrados, com a mesma frequência, no painel de indivíduos normolipidémicos na população portuguesa. Se a presença destes polimorfismos origina uma diminuição dos valores de cLDL a presença destes em indivíduos com diagnóstico clínico de FH pode estar a diminuir os níveis de cLDL, originando, erradamente, um fenótipo mais suave nestes indivíduos.

Um conhecimento aprofundado da função do gene *PCSK9* no ciclo do LDLR é extremamente importante para um melhor conhecimento e compreensão dos distintos fenótipos causados por mutações neste gene, principalmente nas mutações associadas a hipocolesterolemia que originam uma diminuição muito significativa dos valores de cLDL, correspondendo a uma diminuição, até 88%, da ocorrência de um evento cardiovascular (Cohen *et al.* 2006). Com base nestes conhecimentos, a indústria farmacêutica tem desenvolvido vários estudos no sentido de produzir um anticorpo que iniba a interação PCSK9:LDLR de forma a impedir a degradação do LDLR e originando um aumento de recetores na superfície celular. Encontra-se descrito que a percentagem de doentes com diagnóstico clínico de FH e mutações no gene *PCSK9* é inferior a 1% (Abifadel *et al.* 2009; Soutar 2010). No EPHF, foram detetadas alterações no gene *PCSK9*, causadoras de FH, em 0,68% dos CI com diagnóstico clínico, valor similar aos verificados nas restantes populações (Huijgen *et al.* 2010; Marduel *et al.* 2010; Benn *et al.* 2012). Caso a patogenicidade das 2 novas alterações encontradas se confirme, a percentagem de CI com uma mutação no gene *PCSK9* passará para 1,7%, valor acima do descrito em outras populações.

A pesquisa de mutações no gene *PCSK9* é extremamente importante para completar o estudo molecular dos indivíduos referenciados ao EPHF. O estudo deste gene permitiu identificar a causa genética da hipercolesterolemia em 3 indivíduos, bem como identificou 2 novas alterações, em 6 indivíduos, que aguardam os estudos funcionais para comprovação da patogenicidade, podendo contribuir para uma melhoria do conhecimento da interação PCSK9:LDLR. A confirmar-se a patogenicidade das alterações p.Ala62Asp e p.Pro467Ala, o CI que apresenta ambas será o primeiro caso a ser reportado como indivíduo heterozigoto composto para o gene *PCSK9* (Sjouke *et al.* 2014).

## IV.2 Pirosequenciação

Com o avanço das técnicas de sequenciação, é possível sequenciar um gene inteiro de uma forma mais rápida. De modo a identificar a causa genética dos CI com diagnóstico molecular de FH e sem uma mutação identificada num dos três genes estudados no âmbito do EPHF, foram selecionados, em 2009, 65 CI para o estudo de pirosequenciação através do equipamento *454 life science*. O estudo dos CI por pirosequenciação foi realizado pela empresa Biocant em regime de prestação de serviço. A pirosequenciação permite sequenciar vários genes num grande número de indivíduos - por este motivo optou-se por estudar as amostras num *pool*, pois era menos dispendioso e era a técnica mais usada na altura. Esta abordagem teve como desvantagem a necessidade de confirmação de todas as alterações nas 65 amostras por sequenciação de Sanger, de forma a identificar quais as amostras que apresentavam as alterações detetadas.

Uma vez que o gene *LDLR* tinha sido estudado anteriormente no âmbito do EPHF, não era expectável que fossem encontradas alterações neste gene. No entanto, verificou-se que dois dos 65 CI apresentavam uma alteração no exão 6 e 9. A mutação encontrada no exão 6 não foi detetada por falha técnica na análise do eletroforetograma de sequenciação do fragmento em causa. Esta falha foi já colmatada com o desenho de novos *primers* em locais sem a presença de SNPs que originam uma melhor leitura do eletroforetograma. No caso do CI que apresentava a alteração p.Ala431Thr, no exão 9, verificou-se que o *primer* reverso hibridava numa zona onde existia o SNP rs6413505. O CI com a mutação p.Ala431Thr era heterozigótico para este SNP, tendo-se verificado, que a alteração se encontrava no mesmo alelo que o SNP impedindo a sua amplificação. A deteção deste problema levou a que o mesmo fosse corrigido com o desenho de novos *primers* realizando-se posteriormente o estudo de 191 CI com este SNP levando à identificação de 14 CI com alterações no exão 9, aumentando o número de CI identificados no EPHF com uma alteração.

No âmbito do EPHF somente 2 fragmentos dos exões 26 e 29 do gene *APOB* tinham sido analisados, como é prática comum no estudo molecular da FH. A nossa hipótese era que o gene *APOB* poderia conter mais alterações patogénicas do que inicialmente previsto. A análise dos resultados levou à identificação de 32 possíveis alterações neste gene distribuídas por vários exões. No entanto, após a sequenciação de Sanger, apenas 7 destas alterações foram confirmadas, tendo sido ainda detetadas 3 alterações não



reportadas por pirosequenciação. Esta observação comprova a fraca qualidade dos resultados obtidos por este método. No entanto a identificação destas 10 novas alterações no gene *APOB* representam os resultados mais significativos obtidos por esta abordagem.

Apesar de a pirosequenciação ser um método bastante utilizado hoje em dia, à data da realização do estudo (2009) no equipamento *454 life science*, era uma técnica que estava em desenvolvimento com *softwares* de análise ainda limitados que tiveram melhorias significativas com o passar dos anos. A abordagem escolhida (*pool* de amostras) poderá estar na origem da elevada taxa de falsos negativos, assim como da baixa cobertura nos genes estudados (*LDLR* e *APOB*). Existiam na altura algumas limitações associadas à pirosequenciação: várias alterações apresentavam uma estimativa de alelo entre 0,5-0,8 dificultando a interpretação dos resultados; inserções ou deleções de uma base depois de vários nucleótidos idênticos, assim como alterações no início ou final dos fragmentos, podem não ser consideradas e originar um número falso de alelos estimados (Tsialis *et al.* 2010). Estas limitações foram mais tarde ultrapassadas com o desenvolvimento de novos *softwares*.

Em 2012 a empresa Progenika (Espanha) desenvolveu um método para o diagnóstico da FH com base no equipamento *GS Junior system* (pirosequenciação), efetuando a pesquisa de mutações nos genes *LDLR*, *PCSK9*, *LDLRAP1*, e em fragmentos dos exões 26 e 29 onde se encontram as mutações mais comuns no gene *APOB*. De forma a testar esta nova abordagem metodológica, em 2012 foram selecionados 45 CI com diagnóstico clínico de FH para o estudo molecular através deste equipamento. A seleção dos CI teve em consideração o fenótipo agressivo de FH, bem como não apresentar alterações patogénicas num dos três genes estudados no âmbito do EPHF. Uma vez que a primeira abordagem de pirosequenciação originou um resultado tão fraco, foram também enviados 12 controlos com uma mutação identificada anteriormente nos genes *LDLR*, *APOB* e *PCSK9* de forma a validar a técnica. Com este estudo pretendia-se também verificar se o método utilizado no âmbito do EPHF estava a ser realizado em boas condições (ou seja, que a abordagem identificava com sucesso mutações nos genes conhecidos como causadores de FH). Este método revelou uma especificidade e sensibilidade de 100%. tendo sido detetadas todas as alterações enviadas como controlos. Este estudo permitiu também verificar que o método de PCR e sequenciação de Sanger para o diagnóstico de FH realizado no nosso laboratório apresenta uma

especificidade e sensibilidade de 100%. Foi descrito recentemente por um grupo no Reino Unido que o método de diagnóstico por eles utilizado era falível, uma vez que 40% dos indivíduos estudados por métodos de rotina quando foram estudados por sequenciação de exoma apresentavam uma alteração num dos genes associados à FH (Futema *et al.* 2012). Este grupo utilizava diferentes abordagens para o diagnóstico molecular como SSCP, DGGE, *amplification refractory mutation systems* (ARMS) e *High Resolution Melt* (HRM) (Taylor *et al.* 2007). No entanto é sabido que estas abordagens originam um grande número de falsos negativos, devendo ser esta a explicação para a grande percentagem de indivíduos identificados posteriormente por sequenciação de exoma neste grupo.

O gene *LDLRAP1* não tinha sido anteriormente estudado nos 57 CI enviados. Através deste método, foi possível identificar 5 CI com quatro alterações, p.Ile132Val p.Thr218Ile, p.Arg238Trp e p.Asp293His, neste gene. A alteração p.Arg238Trp encontra-se descrita como polimorfismo, sem estudos funcionais descritos (Soutar *et al.* 2003). No entanto, um estudo realizado por Hubacek & Hyatt (Hubacek and Hyatt 2004) avaliou o efeito deste polimorfismo nos valores de CT, em 2 grupos distintos: 100 indivíduos caucasianos com valores de CT acima do percentil 90% (250mg/dL) e 100 indivíduos com um percentil abaixo dos 10% (140mg/dL), tendo verificado não existirem diferenças estatisticamente significativas nas frequências de genótipo ou alelos entre estes grupos. Este estudo sugere, também, que polimorfismos no gene *LDLRAP1* serão determinantes genéticos improváveis nos níveis elevados de cLDL plasmáticos (Hubacek and Hyatt 2004). A alteração p.Thr218Ile encontra-se referenciada na base de dados 1000 Genome com uma frequência de 0,6%, não existindo no entanto dados na literatura sobre estudos funcionais, no entanto esta alteração não co-segrega com a hipercolesterolemia na família. A alteração p.Asp293His não se encontra descrita na literatura nem referenciada em nenhuma base de dados de estudos de *exome sequencing* nem em base de dados de SNP. A análise *in silico* da maioria dos programas bioinformáticos indica que esta não será patogénica, no entanto esta alteração co-segrega na família com a hipercolesterolemia sendo este um bom indicativo da sua patogenicidade. No entanto o estudo funcional da alteração p.Asp293his deveria ser realizado uma vez que a alteração co-segrega com o fenótipo de hipercolesterolemia apresentado pela família do lado paterno. A alteração p.Ile132Val segundo a previsão *in silico* indica que não será patogénica e não co-

segrega na família com a hipercolesterolemia, pelo que nesta família a justificação do fenótipo não deverá estar relacionada com esta alteração.

As alterações foram detetadas em heterozigotia e apenas mutações em homozigotia encontram-se descritas como originando hipercolesterolemia familiar autossómica recessiva, desconhecendo-se assim o risco destes indivíduos desenvolverem uma DCV a longo prazo (Soutar *et al.* 2003), ou se a mutação em heterozigotia pode ser causadora de doença. De forma a verificar a patogenicidade das alterações poder-se-ão realizar estudos funcionais. No entanto, uma vez que a proteína LDLRAP1 não é expressa em todo o tipo de células, a escolha do modelo celular a utilizar no estudo funcional é fundamental para a correta interpretação dos resultados. Uma opção será realizar estudos funcionais a partir de linfócitos dos indivíduos com estas alterações e quantificar a taxa de internalização das partículas de LDL marcadas com um fluoróforo de forma a permitir a sua quantificação por citometria de fluxo, avaliando se alterações em heterozigotia podem originar um fenótipo de hipercolesterolemia. A realização destes estudos não foi possível uma vez que era necessária nova colheita e nenhum dos CI se mostrou disponível.

Com os resultados obtidos pelo *GS Junior system* (utilizando a abordagem desenvolvida especificamente para o diagnóstico molecular da FH), verificou-se uma melhoria significativa em relação aos resultados obtidos anteriormente pelo equipamento *454 life Science*, não se tendo verificado nenhum falso negativo ou positivo nas 12 amostras enviadas para validação do método. Outra vantagem deste método de diagnóstico por pirosequenciação é a possibilidade de deteção dos grandes rearranjos. O diagnóstico da FH é moroso e dispendioso, pelo que vários laboratórios têm procurado desenvolver um diagnóstico que seja eficiente em termos de custos. Este estudo foi realizado no âmbito de um projeto de investigação e o preço aplicado pelas 45 amostras não foi o usualmente aplicado para um diagnóstico molecular de FH. O custo deste serviço é ainda elevado, situando-se por volta dos 600-700€, preço equivalente aos métodos utilizados anteriormente (PCR, Sequenciação de Sanger e MLPA), embora este método requeira menos horas de trabalho por parte do operador. Com o conhecimento e equipamentos atuais o diagnóstico molecular da FH tornar-se-á menos moroso e dispendioso, podendo-se realizar de uma só vez a pesquisa de alterações em todos os genes associados à FH.

Através da abordagem de diagnóstico molecular da FH por pirosequenciação pelo equipamento *GS Junior system* foi possível corrigir-se alguns problemas técnicos, permitindo também identificar novas mutações no gene *APOB* em regiões habitualmente não estudadas, aumentando a identificação de doentes com FH no EPHF. A combinação das técnicas PCR e sequenciação de Sanger revelou ser um método de diagnóstico fiável para a FH. O LIPOnext demonstrou também alta sensibilidade e especificidade com a vantagem de também analisar o gene *LDLRAP1*. O estudo deste gene pode levar a identificação de novas alterações que poderão ser responsáveis pelo do aumento dos níveis de CT e cLDL.

### IV.3 Estudos funcionais

#### IV.3.1 Estudos funcionais de alterações no gene *APOB*

Através da abordagem de pirosequenciação pelo equipamento *454 life science*, foram identificadas 10 possíveis alterações nos exões 19, 22, 24, 26 e 29 do gene *APOB*. De forma a verificar a patogenicidade das alterações foram realizados estudos funcionais em células U937, HepG2 e linfócitos. Uma vez que estes ensaios são realizados com as partículas de LDL dos doentes, não foi possível verificar a patogenicidade de todas as alterações tendo sido apenas estudados os portadores de 4 destas alterações p.Asp1113His, p.Arg1164Thr, p.Tyr1247Cys e p.Gln4494del se mostraram disponíveis para realização de uma nova colheita. Os estudos funcionais das variantes p.Arg1164Thr e p.Gln4494del mostraram existir uma redução de cerca de 40% na ligação da partícula de LDL ao recetor das LDL, assim como na internalização do complexo LDL:LDLR, quando comparados com *wild type*. O mesmo foi observado no ensaio com as células U937, onde se verificou uma redução de 50% na proliferação celular, quando comparadas com *wild type*. Os resultados dos estudos funcionais demonstraram que as alterações originavam uma diminuição da ligação da partícula de LDL ao seu recetor, sendo esta diminuição semelhante ao controlo APOB3527. O estudo funcional das alterações p.Asp1113His e p.Tyr1247Cys revelou que as mesmas não são patogénicas, como esperado, uma vez que se verificou que ambas não co-segregam na família com a hipercolesterolemia e apresentavam uma frequência de 2% num painel de normolipidémicos (Alves *et al.* 2013). Também as alterações p.Asp2213del e p.Ser3279Gly deverão ser polimorfismos, uma vez que estão presentes

com uma frequência de 1-2% num painel de normolipidémicos. Todas as alterações foram verificadas quanto à sua presença num painel de 96 CI com diagnóstico clínico de FH mas sem uma alteração num dos 3 genes associados à FH, tendo-se verificado a presença das alterações p.Asp1113His e p.Ser3279Gly em 5 destes CI (Alves *et al.* 2013). O estudo funcional das restantes 4 alterações estão programados para futuros trabalhos (Figura IV.1).

Neste trabalho foram utilizados 3 tipos de células diferentes de forma a validar os resultados obtidos, tendo sido o ensaio nas células U937 considerado o melhor método por ser o que apresenta menos erros associados a todo o procedimento e por isso maior fiabilidade de resultado (Alves *et al.* 2013). Estudos anteriores desenvolvidos por Van der Broek *et al.* indicam que a sensibilidade e especificidade deste método é de 87,5% e 100%, respetivamente (Van den Broek *et al.* 1994). Os estudos funcionais em alterações no gene *APOB* em linfócitos foram realizados pela primeira vez no âmbito deste trabalho (Alves *et al.* 2013).

A mutação p.Arg1164Thr no gene *APOB* encontrada neste trabalho não se situa na região *consensus* de ligação do ligando ao LDLR, nem na região posteriormente proposta por Krisko e Etchebest (Krisko and Etchebest 2007). No entanto os resultados dos ensaios funcionais demonstraram que esta alteração origina uma diminuição na afinidade da partícula LDL para o seu recetor. Encontra-se descrito na literatura que alterações no gene *APOB* que leva à substituição de em resíduos de arginina causam diminuição da afinidade do complexo LDL:LDLR (Mahley *et al.* 1977; Weisgraber *et al.* 1978; Mahley *et al.* 1979) podendo ser esta a explicação neste caso. Tal como acontece com a mutação mais comum APOB3527, também este aminoácido pode ser importante para a correta conformação da proteína alterando-a e diminuindo a afinidade para o recetor das LDL. O facto de se ter verificado baixa penetrância nos indivíduos mais jovens da família do CI 25023 pode também estar relacionado com a idade, como acontece na dislipidemia familiar combinada (Brouwers *et al.* 2012).

A mutação p.Gln4494del encontra-se na cauda da proteína APOB, numa zona importante para a sua correta conformação e estabilidade à superfície da partícula de LDL. Esta alteração foi encontrada no CI 28120 e alguns dos seus familiares (Figura III.2.10), tendo-se verificado existir uma baixa penetrância em alguns indivíduos. Contudo ao contrário do verificado na família 25023, a baixa penetrância não parece estar relacionada com a idade, podendo dever-se com outros fatores genéticos e/ou

ambientais. Nesta família verificou-se que a avó materna (I:2, Figura III.2.10) do CI tinha um fenótipo de hipercolesterolemia mas não apresentava a alteração no gene *APOB*. Uma vez que existia causa de baixa penetrância e a alteração se situava numa zona onde nunca tinham sido descritas alterações no gene *APOB*, foi sequenciado o exoma do CI, de forma a encontrar a explicação para o fenótipo apresentado nesta família. Identificou-se por esta técnica o polimorfismo p.Asp19His (rs11887534) no gene *ABCG8*, estando este associado a formação de cálculos biliares de colesterol (Buch *et al.* 2007). Os transportadores ABC são proteínas transmembranares que facilitam o transporte de substratos específicos através da membrana celular. Nos eucariotas, os transportadores ABC encontram-se divididos em 7 subgrupos com base na sua sequência e organização (Dean and Allikmets 1995). Entre estes transportadores ABC, encontram-se o *ABCG5* e *ABCG8* que são proteínas que formam heterodímeros funcionais, localizados na membrana apical dos enterócitos e membranas canaliculares dos hepatócitos, agindo como exportadores eficientes de colesterol para a bÍlis (Yu *et al.* 2005). O *ABCG8* encontra-se também no epitélio das vesÍculas biliares e um aumento da sua expressão estÁ associado a litÍases biliares (cÁculos biliares) (Yoon *et al.* 2010). Estas desenvolvem-se quando a bÍlis contém excesso de colesterol e quantidade insuficiente de sais biliares. Segundo Buch *et al.* este polimorfismo (rs11887534) aumenta a eficiêncÍa do transporte de colesterol do fÍgado para a bÍlis ocorrendo saturaçÁo de colesterol e consequentemente cÁculos biliares (Buch *et al.* 2007). Este polimorfismo encontra-se descrito na literatura como estando associado a um aumento da suscetibilidade de desenvolver litÍases biliares de colesterol, bem como um aumento dos nÍveis plasmÁticos de colesterol LDL, sendo esta suscetibilidade maior nas mulheres (Chen *et al.* 2008; Srivastava *et al.* 2010). Os familiares I:2, I:1, II:2, bem como o CI (III.1), apresentam este polimorfismo tendo-se verificado que o colesterol total e LDL é mais elevado nos familiares do sexo feminino (Figura III.2.10). A avó materna (I:2) efetuou uma colecistectomia (exérrese da vesÍcula) há poucos anos, o que reforça a ideia que este polimorfismo possa afetar mais o transporte de colesterol do fÍgado para a bÍlis nas mulheres justificando um aumento dos nÍveis plasmÁticos de colesterol total e LDL (Buch *et al.* 2007; Chen *et al.* 2008; Srivastava *et al.* 2010; Alves *et al.* 2013). O estudo do exoma nesta família, permitiu uma melhor caracterizaçÁo do perfil lipÍdico explicando a falta de co-segregaçÁo do fenótipo com a alteraçÁo no gene no gene *APOB*.

Para melhor compreender a baixa penetrância verificada nas duas famílias, foram também realizados estudos moleculares a todos os exões do gene *PCSK9*, a um fragmento do gene *APOC3* contendo a alteração p.Arg19X, foi caracterizado o genótipo *APOE* de todos os indivíduos para despiste de alterações que causam hipocolesterolemia. Os resultados obtidos não explicam o fenótipo mais suave verificado nestas famílias.

Em 2012 Motazacker *et al* (Motazacker *et al.* 2012) reportaram 2 novas mutações no gene *APOB* (p.Arg3059Cys e p.Lys3367Asn) em doentes com FH clínica que não apresentavam co-segregação completa, observando-se casos de baixa penetrância. Estas alterações no gene *APOB* estão localizadas fora da região *consensus* de ligação do ligando ao recetor das LDL (aminoácidos 3386 ao 3396), mas localizam-se na região postulada por Krisko e Etchebest (Krisko and Etchebest 2007), situação diferente das 2 outras alterações reportadas neste trabalho. Após o estudo funcional em células HepG2, estes autores observaram uma redução de 28% para a mutação p.Arg3059Cys e de 49% para a mutação p.Lys3367Asn na internalização das partículas quando comparados com *wild type* (Motazacker *et al.* 2012). No entanto estes autores apenas realizaram estudos funcionais nos CI e em células HepG2, abordagem diferente da utilizada neste trabalho de doutoramento, onde se realizou o ensaio funcional em 3 tipos de células diferentes de modo a validar os resultados. Além disso, no presente trabalho, foram também, realizados estudos funcionais em familiares o que não é uma abordagem comum, mas importante para reforçar os resultados obtidos.

Como referido anteriormente, a penetrância das alterações nas famílias dos casos índice 25023 e 28120 não era total, ou seja, existiam indivíduos com a alteração mas sem fenótipo de FH (Alves *et al.* 2013). Estes resultados são coerentes com o descrito na literatura, que referem que mutações no gene *APOB* associadas à FH não apresentam uma penetrância de 100%, assim como doentes com mutações no gene *APOB* apresentam, usualmente, um fenótipo moderado da doença (Myant 1993; Vrablík *et al.* 2001). É sabido que mutações no gene *APOB* impedem ou diminuem a ligação do complexo LDL:LDLR. No entanto pode ocorrer um mecanismo compensatório do LDL ao LDLR através da apoE, assim como um aumento na internalização da VLDL remanescente, originando um fenótipo mais suave nos doentes com estas alterações. É sabido que a remoção das lipoproteínas de densidade intermédia (IDL) da corrente sanguínea continua a ser efetuada através da apoE, mas as partículas *small dense* LDL

(sdLDL) continuam em circulação. Sendo estas partículas altamente aterogénicas, apesar dos doentes não apresentarem um fenótipo típico de FH, apresentam um risco aterogénico elevado e consequentemente risco de desenvolver uma DCV prematura (Nigon *et al.* 1991; März *et al.* 1993).

Apenas uma parte dos doentes com mutações no gene *APOB* cumprem os critérios clínicos de FH (Miserez and Keller 1995), mas a não inclusão destes doentes em estudos moleculares podem excluir doentes com mutações no gene *APOB* que apresentam um risco elevado de desenvolverem uma doença cardiovascular. Uma vez que todo o gene da *APOB* não é estudado por rotina, os resultados de Motazacker e os apresentados neste trabalho, indicam que provavelmente a *APOB* deve conter mais mutações patogénicas fora da região consensual que não estão a ser identificadas, já que a prática comum na rotina laboratorial é a análise dos fragmentos que contem as 4 alterações mais comuns, ou seja estudo molecular a um fragmento do exão 26 e outro do exão 29. Com base nestes resultados o estudo de todo o gene *APOB* deveria ser incluído no diagnóstico molecular da FH. Com o desenvolvimento de novas plataformas, será possível incluir a sequenciação de todo o gene *APOB* num diagnóstico molecular de FH através de pirosequenciação, sem aumentar o custo do diagnóstico. A implementação deste método está a ser neste momento desenvolvida no nosso laboratório.

#### **IV.3.2 Estudos funcionais de alterações no gene *LDLR***

No estudo molecular no âmbito do EPHF identificaram-se 2 alterações na região do promotor do gene *LDLR*. A alteração c.-135G>C, umas das mais comuns na população portuguesa, tinha sido descrita anteriormente apresentando estudos funcionais comprovando a patogenicidade, ou seja, a não expressão do alelo com a mutação. A segunda alteração detetada, c.-13A>G, foi descrita pela primeira vez na população portuguesa e publicada posteriormente por Usifo *et al* como possivelmente não patogénica com base na previsão de programas bioinformáticos (Usifo *et al.* 2012). O estudo funcional da alteração c.-13A>G foi por esta razão realizado no âmbito deste trabalho através de RT-PCR a partir de mRNA do doente com a alteração. O resultado funcional negativo foi de encontro ao previsto, uma vez que a alteração não co-segregava completamente na família com a hipercolesterolemia e a previsão *in silico* realizada no âmbito deste trabalho com base em 3 programas bioinformáticos não foi



unânime na classificação da alteração como patogénica. Na região de 200 pb que antecede o codão de iniciação da tradução encontram-se três repetições diretas imperfeitas, duas das quais interatuam com o fator de transcrição Sp1 para promover a transcrição e a terceira contém um local de regulação positiva para esteróis (SER, *sterol regulatory element*), as TATA boxes e um *cluster* de locais de iniciação do mRNA (Goldstein 95 e Sato e Takano 95). A alteração c.-135G>C localiza-se na primeira repetição Sp1 que é importante para a interação com o fator de transcrição e por este motivo uma alteração neste local impede a transcrição correta levando a não expressão do alelo com a mutação. A alteração c.-13A>G não se situa em nenhuma das regiões importantes na região do promotor, não devendo influenciar a expressão normal do gene, embora já tenham sido descritas outras alterações nesta zona nenhuma das quais com estudos funcionais comprovando a patogenicidade das alterações (Day *et al.* 1997; Fouchier *et al.* 2005). Os estudos funcionais são bastante importantes para comprovar a patogenicidade das alterações de forma a se obter um melhor conhecimento da doença e dos genes responsáveis.

No gene *LDLR*, foram encontradas 19 alterações de *splicing* das quais 5 tinham estudos funcionais realizados por outros grupos comprovando a patogenicidade das mutações (Hobbs *et al.* 1992; Leren *et al.* 1994; Lombardi *et al.* 1995; Mak *et al.* 1998). O estudo funcional a 12 alterações verificou que 8 destas originavam *skipping* de um exão ou retenção de parte ou totalidade do intrão, levando à produção de uma proteína truncada. Porém, 3 destas alterações não influenciavam o correto processamento do mRNA. A alteração p.Val797Met, apesar de ter sido classificada como mutação de *splicing*, é também uma mutação pontual. No entanto, o nucleótido alterado é o último do exão 16 e por esta razão importante para o correto processamento do mRNA. A remoção dos intrões e a ligação dos exões é realizado por meio de ribonucleoproteínas que reconhecem as extremidades típicas dos intrões, 5'-GU e AG-3'. Se alguma destas extremidades se encontrar alterada as ribonucleoproteínas não reconhecem os locais de remoção e ligação ocorrendo deleção de exões ou retenção de intrões (Lehrman *et al.* 1987; Dogan *et al.* 2007). Todas as alterações estudadas foram amplificadas a partir de mRNA dos doentes por RT-PCR, visualizadas em gel de agarose, mas também sequenciadas de modo a verificar o real efeito das alterações, uma vez que algumas alterações originam a retenção de apenas 2 nucleótidos que são impossíveis de visualizar num gel de agarose.

Todas as alterações de *splicing* foram verificadas, quanto à sua presença, num painel de normolipidémicos tendo-se verificado a presença das alterações c.1061-8T>C e 2140+5G>A reforçando o resultado obtido *in vitro* de não patogénicas. As alterações de *splicing* encontradas no decorrer deste trabalho não foram estudadas por PCR quantitativo em tempo real (RT-qPCR) não sendo possível saber qual a percentagem de transcritos que tem um correto processamento do mRNA. Esta caracterização é importante, uma vez que os doentes com alterações de *splicing* apresentavam por vezes um fenótipo moderado e muito variável. Uma das justificações poderá ser o facto de o alelo normal ser mais expresso que o alelo mutado, havendo assim níveis quase normais da proteína. Os estudos por RT-qPCR estão programados para estudos futuros sendo de grande importância, uma vez que podem alterar a classificação dos CI que apresentam estas mutações.

Foram detetadas 8 alterações sinónimas no decorrer deste trabalho, usualmente este tipo de alterações não originam qualquer efeito na proteína, mas podem afetar o correto processamento do mRNA e levar à produção de proteínas truncadas, como descrito anteriormente por Bourbon *et al.* e Defesche *et al.* que verificaram que alterações sinónimas situadas nos exões 9 e 4 respetivamente, originam este tipo de efeito na proteína (Bourbon *et al.* 2007; Defesche *et al.* 2008). À luz deste conhecimento todos os grupos que estudam FH deverão reportar e estudar funcionalmente este tipo de alterações. Foram estudadas funcionalmente 3 alterações sinónimas, tendo-se verificado que nenhuma afetava o normal processamento do mRNA. O estudo funcional de todas as alterações encontradas, mesmo as sinónimas, deve ser efetuado de modo a caracterizar e classificar corretamente as alterações, bem como a diagnosticar corretamente o doente. Os estudos funcionais de *splicing* descritos neste trabalho são razoavelmente fáceis de realizar, embora limitados pela disponibilidade de uma nova amostra de sangue do CI.

No total foram identificados 29 CI com uma mutação de *splicing* representado cerca de 10% dos CI estudados no EPHF, sendo este resultado inferior ao verificado noutras populações, principalmente no norte da Europa onde a percentagem deste tipo de alterações é de aproximadamente 23% (Dedoussis *et al.* 2004), mas semelhante ao verificado em França e Espanha (13% e 9,8%) (Marduel *et al.* 2010; Palacios *et al.* 2012), bem como nas bases de dados internacionais para mutações reportadas nos genes associados à FH (<http://www.ucl.ac.uk/ldlr/> e <https://grenada.lumc.nl/LOVD2/UCL->

Heart/). No entanto, algumas destas alterações são reportadas à base de dados sem os estudos funcionais realizados, o que pode alterar a percentagem de mutações de *splicing* a nível mundial.

No EPHF foram encontradas 59 alterações pontuais no gene *LDLR*, das quais apenas 22 alterações tinham estudos funcionais comprovando a sua patogenicidade. No âmbito deste trabalho foi realizado o estudo funcional a 4 das 59 alterações (p.Gly76Trp, p.Ile441Thr, p.Gly545Trp e p.Cys698Phe) de forma a verificar a patogenicidade das mesmas (Figura IV.1). Os ensaios foram realizados por citometria de fluxo a partir de linfócitos dos doentes, motivo pelo qual apenas se realizou o estudo a 4 alterações diferentes, uma vez que este método depende fortemente da obtenção de nova amostra. A alteração p.Gly76Trp foi detetada até agora apenas da população portuguesa, encontra-se no exão 3 do gene *LDLR* no domínio de ligação ao ligando, tendo alterações neste exão sido descritas como causadoras de doença (Rubinsztein *et al.* 1993; Bertolini *et al.* 1999). Apesar da análise *in silico* apontar para que a alteração fosse patogénica e não se tendo verificado a presença da alteração num painel de normolipidémicos, o estudo de co-segregação da alteração na família revelou que alguns familiares normolipidémicos apresentavam a alteração, levando a questionar a sua patogenicidade. Os estudos funcionais para a alteração p.Gly76Trp mostraram existir apenas uma ligeira redução na ligação do complexo LDL:LDLR ( $\approx 6\%$ ), bem como na internalização do complexo em comparação ao *LDLR* wt ( $\approx 3\%$ ), sendo os níveis de expressão normais. Anteriormente foi descrita no EPHF, a alteração p.Gly269Asp, (Bourbon *et al.* 2008) que foi considerada patogénica pela análise *in silico*, mas onde não se observava co-segregação na família. Etxebarria *et al.* confirmaram, anos mais tarde, através de estudos funcionais que a alteração p.Gly269Asp não era patogénica (Etxebarria *et al.* 2012). A verificação da co-segregação da alteração com a hipercolesterolemia na família parece ser um dos melhores indicadores na avaliação *in silico* da patogenicidade das alterações, tendo sido uma referência em todas as alterações encontradas no decorrer deste trabalho.

A análise *in silico* classificou as alterações p.Gly545Trp e p.Ile441Thr como provavelmente patogénicas tendo-se verificado também, que ambas co-segregam com o fenótipo de hipercolesterolemia nas famílias. Estas alterações estão localizadas no domínio de homologia ao precursor do EGF e foram encontradas em 8 famílias do EPHF. Alterações neste domínio podem impedir o normal posicionamento do LDLR na

superfície impedindo a ligação ao ligando, bem como, impedir a dissociação das lipoproteínas do recetor no endossoma (Davis *et al.* 1987). Os estudos funcionais às alterações demonstraram uma redução de 5-15% na ligação do complexo LDL:LRLR e de 35-40% na internalização do complexo LDL:LRLR. Apesar de ambas terem sido classificadas como patogénicas neste trabalho, os resultados obtidos são na nossa opinião *bordeline*, especialmente para a alteração p.Ile441Thr, devendo ser realizados novos estudos funcionais em sistemas de expressão heteróloga.

A alteração p.Cys698Phe encontra-se no exão 14 do *LDLR* na repetição C do domínio de homologia com o precursor do EGF (Hobbs *et al.* 1986). Esta região do *LDLR* é altamente conservada entre espécies e as primeiras mutações pontuais descritas funcionalmente no *LDLR* foram em resíduos próximos (Knight *et al.* 1989). O estudo de co-segregação da alteração nas famílias com a hipercolesterolemia verificou que um indivíduo hipercolesterolémico não apresentava a alteração e a análise *in silico* classificou a alteração como provavelmente patogénica. Os estudos funcionais mostraram existir uma redução de cerca de 50% na ligação e internalização, mas uma expressão normal. A alteração parece originar um atraso na maturação do LDLR bem como uma menor afinidade para o ligando, uma vez que impede que a  $\beta$ -hélice remova o chaperone necessário no processo de maturação no retículo endoplasmático (Mark *et al.* 2011) tendo sido classificada como patogénica. No familiar hipercolesterolémico que não apresentava a alteração verificou-se que os níveis de triglicéridos eram elevados, podendo a origem do fenótipo neste indivíduo ser poligénica ou ambiental.

Os primeiros estudos funcionais no *LDLR* foram desenvolvidos por Brown e Goldstein, e baseavam-se na determinação dos níveis de ligação, internalização e degradação de LDL marcado radioactivamente com  $^{125}\text{I}$  na porção proteica do LDL e eram realizados em fibroblastos e linfoblastos de doentes (Goldstein and Brown 1979; Knight *et al.* 1987; Hobbs *et al.* 1992). Foram também desenvolvidos estudos com base em sistemas de expressão heteróloga *in vitro* que permitem estudar o efeito do alelo alterado na função do LDLR, sem interferência do alelo normal ou da combinação de outros fatores genéticos nos doentes (Chang *et al.* 2003; Silva *et al.* 2012). Posteriormente e na tentativa de simplificar o diagnóstico molecular de indivíduos com diagnóstico clínico de FH, Raungaard *et al.* (Raungaard *et al.* 1998) desenvolveram um método por citometria de fluxo em linfócitos de indivíduos com suspeita clínica de FH pretendendo que este método fosse informativo quanto ao diagnóstico molecular destes indivíduos,

tendo concluído que não era um bom método de diagnóstico molecular mas que seria um bom método de caracterização das mutações encontradas em indivíduos com FH (Raungaard *et al.* 1998).

Neste trabalho o método utilizado para a realização de estudos funcionais foi a citometria de fluxo para determinar a taxa de ligação e internalização de LDL marcado com 3,3'-dioctadecilindocarbocianina (DiI), um fluoróforo lipofílico, ou com recurso a anticorpos específicos para o LDLR, marcados secundariamente com fluorescência, de forma a avaliar os níveis de expressão do recetor à superfície (Jensen *et al.* 1994; Etxebarria *et al.* 2012). Com a utilização deste tipo de marcação, é também possível analisar a expressão do recetor por microscopia de fluorescência. Apesar de ser um método fiável (Raungaard *et al.* 1998; Raungaard *et al.* 1999; Etxebarria *et al.* 2012) a realização de estudos funcionais em sistemas de expressão heteróloga é o mais recomendado, embora mais dispendioso, tendo a vantagem de não ser necessária a colaboração dos indivíduos com as alterações. As alterações apresentadas neste trabalho irão ser analisadas futuramente neste tipo de sistemas heterólogos, de forma a conseguir uma melhor caracterização, uma vez que alguns resultados obtidos são *borderline* deixando algumas dúvidas quanto à sua interpretação

#### IV.4 Análise *in vitro* vs *in silico*

Encontram-se descritas mais de 1300 alterações no gene *LDLR* como causadoras de FH (Usifo *et al.* 2012), dado que aproximadamente 55% são alterações *missense* e apenas uma parte destas alterações apresentam estudos funcionais comprovando a sua patogenicidade (Leigh *et al.* 2008; Usifo *et al.* 2012). Os estudos funcionais são importantes para a comprovação da patogenicidade das alterações, uma vez que a previsão *in silico* não é suficiente para se aferir sobre a patogenicidade das mesmas e deve ser utilizada com o máximo cuidado. Para um indivíduo com diagnóstico clínico de FH é extramente importante identificar a causa genética da hipercolesterolemia, uma vez que quando identificada a mutação como causadora de FH é necessário a introdução adequada da terapêutica farmacológica, usualmente uma estatina potente ou a combinação de terapêuticas de modo a que os doentes atinjam os alvos terapêuticos desejados. A identificação precoce de crianças com FH é também uma grande vantagem uma vez que estas crianças irão ser acompanhadas por um pediatra que irá implementar

uma dieta rígida e implementação de terapêutica farmacológica adequada, usualmente a partir dos 8 anos de forma a reduzir o risco de desenvolverem uma doença cardiovascular prematura (DeMott *et al.* 2008; NICE 2008; Nordestgaard *et al.* 2013). Na impossibilidade da realização de estudos funcionais a todas as alterações encontradas no EPHF, foi elaborada uma classificação para as alterações cuja patogenicidade era desconhecida de modo a avaliar o seu significado clínico (Figura IV.1). De forma a validar a classificação proposta neste trabalho, foi realizada uma comparação *in silico* vs *in vitro* tendo-se verificado que a previsão *in silico* classificou corretamente 81,5% das alterações *missense* e 100% das alterações de *splicing*. No ano de 2012 os responsáveis pela base de dados internacional de mutações na FH (Usifo *et al.* 2012) realizaram um estudo semelhante com base em 10 programas bioinformáticos, classificando 223 alterações reportadas à base de dados, tendo verificado que 79% destas alterações eram possivelmente a causa da doença. No entanto, alguns dos programas bioinformáticos utilizados por estes autores já não se encontram disponíveis *online*, impedindo a aplicação dos mesmos critérios utilizados por estes autores.

A análise *in silico* para as alterações *missense* reportadas neste trabalho foi realizada com base em 6 programas bioinformáticos que originam uma previsão do efeito da alteração na proteína, tendo por base a alteração do nucleótido, aminoácido, bem como a conservação entre espécies. O efeito do correto processamento do pré-mRNA para as alterações de *splicing* foi analisado em 3 programas bioinformáticos. Após a previsão *in silico* das alterações sem estudos funcionais classificou-se 36 alterações: 19 prováveis, 7 possíveis e 10 neutras.

A análise *in silico* origina uma previsão do efeito que a alteração originará na proteína, contudo os programas bioinformáticos não conseguem prever com exatidão o efeito real que a alteração causará na proteína. Para algumas proteínas, como é o caso da APOB o resultado *in silico* deve ser interpretado ainda com mais atenção, uma vez que esta proteína nunca se encontra livre na circulação sendo difícil obter-se a sua estrutura cristalina (Prassl and Laggner 2009). O avanço tecnológico permitiu melhorar significativamente os programas de previsão bioinformáticos, desenvolvendo novos programas com abordagens de análises diferentes originando uma melhor previsão do efeito da alteração na proteína. Como descrito neste trabalho, após a análise nos programas bioinformáticos, a alteração p.Gly76Trp foi classificada como patogénica, tendo o estudo funcional comprovado o contrário. Neste caso e sempre que possível o

melhor indicador é a observação da co-segregação da alteração com a hipercolesterolemia na família. Contudo a obtenção de familiares nem sempre é um processo fácil o que dificulta este tipo de análise.

Encontra-se descrita na literatura que a prevalência de FH é 1 em 500 na maioria das populações. No entanto foi reportado recentemente que esta prevalência poderia ser de 1 em 200 na Holanda e Noruega (Nordestgaard *et al.* 2013). Esta prevalência foi calculada com base nos indivíduos identificados nestes países, onde existe a maior taxa de indivíduos identificados com FH em todo o mundo. Nestes países foi desenvolvido, em conjunto com os ministérios da saúde de cada país, um estudo nacional para a identificação clínica e molecular destes doentes bem como participação do respetivo tratamento e aconselhamento de CI e familiares com FH (Fouchier *et al.* 2001; Nordestgaard *et al.* 2013). No entanto, nestas populações, uma grande parte das alterações reportadas não tem estudos funcionais que comprovem a sua patogenicidade (Fouchier *et al.* 2005; Huijgen *et al.* 2010; Usifo *et al.* 2012). A não realização do estudo funcional pode estar na origem de uma prevalência diferente da verificada na maioria das populações.

Na Holanda, devido ao grande número de indivíduos identificados (73%), foi desenvolvida uma abordagem baseada em critérios de classificação para alterações com patogenicidade desconhecida, com base na co-segregação das alterações nas famílias, valores de cLDL antes e após iniciação da terapêutica farmacológica, bem como a resposta à mesma e na presença de DCV nos indivíduos com FH (Huijgen *et al.* 2010; Huijgen *et al.* 2012). Contudo, este tipo de abordagem apenas funciona em *cohorts* onde existe um grande número de indivíduos estudados e identificados na mesma família. Uma das maiores limitações encontradas durante este trabalho de doutoramento foi a angariação de familiares para uma informação completa da história clínica da família tendo sido complicado, na maioria dos casos, realizar uma avaliação clara e objetiva quanto à patogenicidade das alterações. A análise *in silico* deve ser considerada informativa e utilizada como ferramenta para melhorar classificar as alterações mas não deverá nunca ser conclusiva.

#### IV.5 Caracterização bioquímica - genótipo vs fenótipo

A maioria dos grupos que se dedica ao estudo da FH avalia e classifica os indivíduos pelo diagnóstico clínico com base nos valores de colesterol total e cLDL dos CI e seus familiares, bem como na ocorrência de DCV nas famílias (Marks *et al.* 2003; Nordestgaard *et al.* 2013). Estes grupos baseiam o diagnóstico clínico em critérios conhecidos e descritos com base em estudos com doentes com FH identificada molecularmente (Group 1991; WHO 1998; DeMott *et al.* 2008) realizando o diagnóstico molecular normalmente apenas nos indivíduos que cumprem estes critérios. O critério clínico utilizado neste trabalho foi adaptado do “Simon Broome Heart Research Trust” (Group 1991) - este apresenta valores de diagnóstico para adultos e crianças classificando os indivíduos com possível FH e FH definitiva (para os indivíduos que apresentam xantomas tendinosos ou uma mutação identificada molecularmente) existindo mais 2 critérios clínicos (Make Early Diagnosis to Prevent Early Deaths- MED PED; e Dutch Lipid Clinic Network criteria - DLCN) que se baseiam num *score* e classificam os indivíduos com FH possível, provável e definitiva (WHO 1998; DeMott *et al.* 2008). No entanto o DLCN não apresenta um *score* para crianças, sendo uma das suas limitações.

Todos os indivíduos referenciados ao EPHF e analisados neste trabalho, apresentam um diagnóstico clínico com base no critério “Simon Broome Heart Research Trust”, tendo-se verificado que em apenas 28% dos CI foi identificada uma mutação patogénica (com estudos funcionais) causadora de FH. No entanto, considerando as alterações patogénicas e as prováveis (classificação *in silico*), a percentagem aumenta para 33% e se considerarmos os CI com qualquer alteração encontrada obtemos uma percentagem de 45%. Usualmente, os grupos que estudam a FH realizam análises estatísticas para comparação genótipo vs fenótipo com todos os CI que apresentam uma alteração, independentemente de terem sido realizados estudos funcionais. Neste trabalho a análise às características bioquímicas dos CI do EPHF foi realizada com base no diagnóstico molecular (diagnóstico definitivo – DD, diagnóstico provável - DP, diagnóstico possível – DPos; com uma alteração neutra (DN) e sem alteração (FH)), por ser a abordagem mais correta a nível funcional.

A análise estatística foi realizada tendo em conta os valores bioquímicos de todos os CI e familiares sem medicação, tendo este facto originado uma diminuição do número de CI e familiares considerados para análise, nomeadamente nos parâmetros do perfil



lipídico específico e diferenciado, tendo sido impossível realizar algumas análises. Nos casos onde foi realizada a análise estatística os resultados obtidos para o perfil diferenciado não deverão ser considerados conclusivos.

Uma vez que as crianças apresentam características clínicas e um perfil lipídico diferente dos adultos, a análise foi realizada em separado sendo a discussão dos resultados apresentada em separado.

#### IV.5.1 Grupo Pediátrico

Na análise estatística realizada ao grupo pediátrico verificou-se não existir diferença estatisticamente significativa quando comparados os valores de CT dos CI com diagnóstico definitivo e provável. O mesmo foi verificado para os valores de apoB, apoA-I e razão apoB:apoA-I. No entanto verificou-se existirem diferenças estatisticamente significativas ( $p=0,04$ ) para os valores de cLDL nas crianças com diagnóstico definitivo e provável, embora este valor seja *borderline*. No entanto para as sdLDL o mesmo não se verificou. As propriedades físico-químicas das sdLDL conferem-lhes um potencial aterogénico elevado - devido ao seu tamanho as sdLDL possuem um fácil acesso ao espaço subendotelial da parede arterial exibindo uma grande afinidade para os proteoglicanos da camada íntima. As sdLDL mostram também uma grande suscetibilidade para a oxidação e captação por macrófagos, o que facilita a formação das células esponjosas promovendo a disfunção endotelial (Gazi *et al.* 2007). O tempo em circulação das partículas de sdLDL é superior ao tempo em circulação das LDL de tamanho médio, pois há um menor reconhecimento das sdLDL pelo LDLR, o que potencia a sua oxidação e deposição na parede arterial (Gazi *et al.* 2007). A avaliação dos níveis de sdLDL é importante pois dá uma informação adicional sobre a quantidade de sdLDL existente em circulação sabendo-se que estes valores estão diretamente relacionado com o risco cardiovascular. Valores elevados de sdLDL podem ser indicativos que estas crianças apresentam uma alteração causadora de doença. Outro dado também observado é que os valores de apoA-I e apoA-II, bem como os de cHDL são superiores nas crianças que apresentam DPos, DN e FH em comparação com DD e DP, verificando-se mesmo que as crianças com Nt são as que apresentam os valores mais elevados. Uma vez que num fenótipo típico de FH os valores de cHDL são normalmente baixos, a causa da hipercolesterolemia nestas crianças deverá ter outra origem. Apesar dos níveis de triglicéridos nestas crianças serem normais, não deverá ser

descartada a hipótese destas terem uma dislipidemia familiar combinada. Todas estas observações fundamentaram a opção de se agregar as crianças com diagnóstico definitivo e provável para posterior análise de correlação fenótipo/genótipo.

Os resultados verificados para as crianças em estudo neste trabalho no que diz respeito aos valores de apoB, apoA-I e razão apoB:apoA-I encontra-se em consonância com os resultados obtidos recentemente no nosso grupo, onde se verificou que o valor de apoB se mantém constante nos doentes com FH independentemente do tipo de mutação que apresentem, e de se tratar de CI ou familiares (Medeiros *et al.*, 2014). A apoB parece ser um bom biomarcador, assim como a razão apoB:apoA-I, tendo sido proposto pelo nosso grupo que para as crianças os critérios clínicos do *Simom Bromme* deveriam ser aplicados em combinação com o *cutoff* de razão apoB:apoA-I  $\geq 0,68$ , por ser o critério que apresentou melhor sensibilidade e especificidade (Medeiros *et al.*, 2014). Estes resultados levam a crer que as crianças com diagnóstico provável apresentam alterações causadoras de doença, uma vez que estas crianças apresentam uma razão apoB:apoA-I de 0,85.

Na análise à correlação fenótipo/genótipo, verificou-se não existirem diferenças estatisticamente significativas por tipo de mutação, ou seja, independentemente do tipo de mutações os valores bioquímicos são semelhantes. A justificação para esta observação pode estar relacionado com o facto das crianças apresentarem um perfil bioquímico com menos influência dos estilos de vida, no entanto, este normalmente altera-se na idade adulto.

Encontra-se descrito na literatura que familiares dos CI com uma mutação causadora de FH apresentam valores de CT e cLDL inferiores (alguns não apresentam mesmo valores de diagnóstico clínico) aos verificados nos CI (Sun *et al.* 1998; Alves *et al.* 2010). Na análise estatística realizada às crianças estudadas neste trabalho verificou-se que os familiares apresentam valores de CT e cLDL cerca de 12% inferiores aos dos CI, estando de acordo com o descrito anteriormente (Alves *et al.* 2010).

As crianças com diagnóstico clínico de FH, independentemente da origem da sua dislipidemia, apresentam vários fatores de risco cardiovascular, para além da hipercolesterolemia - 40% das crianças referenciadas ao EPHF apresentam 2 fatores de risco cardiovascular. Este resultado alerta para a prevenção cardiovascular desde a infância e a importância da aplicação de medidas preventivas e corretivas em idade jovem de forma a reduzir o risco de DCV na idade adulta (Daniels *et al.* 2011). O

aumento da atividade física deve ser uma prioridade uma vez que se verificou que a maioria das crianças faz menos de 3 horas de exercício físico por semana. A correta identificação e estratificação das crianças em risco para a doença cardiovascular na infância, juntamente com um aumento de iniciativas educacionais, implementação de hábitos de vida saudáveis, bem como implementação precoce de terapêutica hipolipemiante nas crianças com FH (Hopkins *et al.* 2011), irá certamente diminuir a percentagem de doença cardiovascular na população adulta.

A análise bioquímica das crianças com diagnóstico clínico de FH levou à identificação de novos biomarcadores para diagnóstico que foram implementados no nosso laboratório, levando a um aumento da taxa de deteção de identificação genética.

As crianças com diferentes tipos de mutação não apresentaram um fenótipo distinto mas crianças com mutações nulas devem estar sobre uma maior vigilância clínica uma vez que só têm 50% de recetores funcionais.

#### **IV.5.2 Grupo Adultos**

A análise estatística ao grupo adultos verificou existirem diferenças estatisticamente significativas para os valores de CT e cLDL ( $p < 0,001$ ) nos CI com diagnóstico definitivo e provável quando comparados com os restantes diagnósticos. Ao contrário do verificado no grupo pediátrico, nos adultos não existem diferenças estatisticamente significativas para os valores de cLDL nos indivíduos com diagnóstico definitivo e provável. Apesar do fenótipo das crianças apresentar menos influência dos estilos de vida, o resultado verificado para os adultos mostra que há uma similaridade nos valores de cLDL para os indivíduos com diagnóstico definitivo e provável. Verificou-se também que os valores de apoB e razão apoB:apoA-I são mais elevados nos adultos com diagnóstico definitivo e provável em comparação com os restantes diagnósticos, uma vez que no grupo adultos os CI com diagnóstico possível apresentam valores mais próximo dos CI com diagnóstico definitivo e provável do que os indivíduos com Nt e SA, resultado que não foi verificado no grupo pediátrico. Uma vez que a análise estatística é realizada com base nos valores obtidos antes do início da terapêutica hipolipemiante, não foi possível realizar a análise estatística para os diferentes diagnósticos moleculares para os valores de sdLDL no grupo adultos. A análise estatística foi realizada apenas para os adultos com diagnóstico definitivo e provável vs

sem alterações, tendo-se verificado que os valores de sdLDL são estatisticamente significativos ( $p=0,01$ ).

Assim como no grupo pediátrico, a análise por tipo de mutação aos CI adultos foi realizada considerando-se apenas os indivíduos com diagnóstico definitivo e provável. Verificou-se que os CI com alterações *missense* apresentavam valores de CT e cLDL estatisticamente significativos mais baixos ( $p=0,03$ ) quando comparados com mutações nulas e de *splicing*, encontrando-se este resultado em concordância com o descrito na literatura - mutações *missense* originam um fenótipo mais suave de FH quando comparado com outros tipos de mutações (Soutar and Naoumova 2007). Os CI com mutações no gene *APOB* apresentam valores de CT e cLDL inferiores aos CI com os restantes diagnósticos, estando este resultado de acordo com o verificado noutros grupos e descrito anteriormente no nosso grupo (Miserez and Keller 1995; Soutar and Naoumova 2007; Alves *et al.* 2013).

Também para o grupo adultos foi realizada uma análise de comparação dos valores CT e cLDL entre familiares e CI com alterações causadoras de FH. Observou-se que os CI apresentam valores superiores de CT e cLDL, 13% e 17% respetivamente, verificando que esta diferença é mais significativa do que no grupo pediátrico. Curiosamente no grupo adultos verifica-se um aumento nos níveis de apoA-I nos CI em relação aos familiares, mas o mesmo não se verifica com os valores de cHDL. Uma vez que a partícula de HDL é maioritariamente constituída por apoA-I seria expectável que os valores de ambos os parâmetros estivessem em consonância. No entanto é importante referir que o número de indivíduos analisados para os valores de apoA-I é inferior, podendo ser esta a justificação para a diferença de valores.

Uma característica clínica da FH é a ocorrência de um evento cardiovascular prematuro, usualmente enfarte agudo do miocárdio, no CI ou num familiar de primeiro grau. Dos 392 CI adultos referenciados ao EPHF, 31% tinham já sofrido um evento cardiovascular uma vez que em apenas 9,5% deste CI foi encontrada uma mutação patogénica. 6,5% dos CI apresentam ainda uma alteração cuja patogenicidade não está comprovada, podendo aumentar o número de CI com DCV e uma mutação causadora de FH. No entanto, o facto de um individuo apresentar uma DCVp não é indicativo de ter FH; na realidade apenas cerca de 20% dos doentes com DCV prematura têm FH (Hopkins 2010).

Os resultados apresentados neste trabalho encontram-se em concordância com os verificados noutras populações - a maioria dos indivíduos sem mutação apresenta valores inferiores aos verificados nos indivíduos com diagnóstico molecular (Talmud *et al.* 2013). Nestes CI a origem da dislipidemia pode ser poligénica ou a alteração encontrar-se num gene do metabolismo lipídico ainda não associada à FH, uma vez que apresentam valores de cLDL elevados. Foi descrito recentemente que a mutação p.Leu167del no gene *APOE* pode originar um fenótipo de FH ou FCHL (Okorodudu *et al.* 2013), sendo esta alteração mais comum em indivíduos com diagnóstico clínico de FH. Em trabalhos futuros, será realizada a pesquisa desta mutação para todos os CI sem mutação encontrada.

A dislipidemia familiar combinada (FCHL) é uma doença poligénica com uma frequência de 1 em 250, verificando fenótipos muito variados nas famílias e que variam ao longo do tempo. Estes indivíduos podem apresentar valores elevados de CT e triglicéridos, tendo também usualmente um valor elevado de apoB. Verificou-se que nos CI do EPHF com uma mutação num dos 3 genes associados à FH 2,7% apresentam valores de triglicéridos superiores a 250mg/dL (o valor máximo de referência para os triglicéridos segundo a ESC é de 150mg/dL). Teoricamente nos indivíduos que apresentam valores de triglicéridos não deveria ser realizado o estudo molecular de FH, uma vez que a origem genética poderia ser a FCHL. Como verificado neste trabalho e em trabalhos anteriores, este facto não é totalmente verdadeiro (Medeiros *et al.* 2010) - o valor dos triglicéridos pode aumentar em indivíduos com hábitos etanólicos excessivos ou com uma alimentação rica em gorduras (Barson *et al.* 2009). Por este facto foram incluídos no diagnóstico molecular do EPHF 43 indivíduos com valores de triglicéridos superiores a 250 mg/dL, tendo-se verificado que 35 não apresentam nenhuma alteração. Nestes CI a causa da dislipidémia poderá estar em genes associados à FCHL ou origem ambiental.

A FH está sub-diagnosticada na nossa população como noutras populações e os CI identificados encontram-se longe de atingir os alvos terapêuticos (Bourbon and Rato 2006; Bourbon *et al.* 2008; Medeiros *et al.* 2010; Nordestgaard *et al.* 2013). Os clínicos devem ter atenção às recomendações europeias e medicar estes indivíduos agressivamente, para se obter uma prevenção cardiovascular efetiva. É necessário identificar indivíduos com FH o mais precocemente possível para que possam ser tratados e aconselhados adequadamente desde jovens, de forma a se reduzir o seu risco

cardiovascular. É mais fácil educar uma criança para a adoção de estilos de vida saudáveis e alertar para a importância da toma regular da terapêutica, do que tentar alterar os hábitos de vida de um adulto.

A identificação da causa genética pode influenciar o prognóstico do indivíduo com FH, uma vez que permite estratificar melhor o risco cardiovascular. O aumento da identificação precoce destes indivíduos pode ter impacto na mortalidade e morbilidade cardiovascular em Portugal. Médicos, enfermeiros, investigadores, decisores políticos e população em geral devem estar atentos para este importante problema de saúde pública. A FH está associada a um elevado risco cardiovascular, mas este risco é modificável se estes indivíduos forem identificados, aconselhados e tratados precocemente.

Neste trabalho a análise estatística foi realizada tendo por base diferentes diagnósticos moleculares que foram obtidos através da classificação das alterações com e sem estudos funcionais. Este tipo de análise não foi realizado anteriormente e permitiu uma melhor caracterização de todos os casos índice referenciados ao EPHF consoante o tipo de alteração que possuem, mostrando que os CI com diagnóstico provável apresentam um fenótipo mais semelhante aos CI com diagnóstico definitivo e que os CI com diagnóstico possível e neutro apresentam um fenótipo mais próximo dos CI sem mutações.

A comparação dos fenótipos apresentados por crianças e adultos com diferentes tipos de mutação revelou que embora em idade pediátrica o fenótipo não dependa do tipo de mutação, na idade adulta já existe uma clara diferença de fenótipos entre portadores de mutações nulas e *missense*. Esta observação reforça a importância do diagnóstico molecular precoce de modo a estratificar o risco cardiovascular no grupo pediátrico, uma vez que as crianças com mutações nulas apresentam um maior risco cardiovascular, necessitando por isso de um aconselhamento especializado e implementação precoce de terapêuticas hipolipemiantes.

#### **IV.5.3 FH homozigótica e heterozigótica composta**

Nos CI homozigóticos e heterozigóticos compostos, foi também pesquisada a presença de mutações no gene *PCSK9* associados a hipocolesterolemia, não tendo sido detetada nenhuma mutação. A opção de se pesquisar alterações associadas à hipocolesterolemia

neste CI prendeu-se com o facto de estes indivíduos não apresentarem um fenótipo típico de homozigóticos.

Para os 5 CI heterozigóticos compostos e 3 CI homozigóticos, com alterações no gene *LDLR*, aplicou-se o mesmo critério de classificação das alterações, tendo-se verificado que 2 CI foram classificados como homozigóticos e apenas 1 CI foi classificado como heterozigótico composto. Os restantes CI foram classificados como heterozigóticos. Talvez por este facto, estes indivíduos não apresentem um fenótipo típico de HoFH. A prevalência de indivíduos com FH homozigóticos é 1 em 1 000 000 o que corresponderia a 10 CI em Portugal com base na classificação aplicada neste trabalho foram identificados, até ao momento, apenas 3 CI (WHO 1998; Bourbon and Rato 2006).

#### IV.6 Sequenciação de exoma

Apesar da FH ser uma das doenças monogénicas melhor estudada, existe num entanto um grande número de doentes com diagnóstico clínico de FH sem mutações num dos genes associados à FH (Kusters *et al.* 2011; Futema *et al.* 2012). Neste trabalho apenas foi identificada uma mutação comprovadamente patogénica em 28% dos CI, sendo necessário nos restantes 108 CI averiguar quanto à patogenicidade da alteração. De forma a clarificar a causa genética da hipercolesterolemia foi realizado sequenciação de exoma a 5 CI, tendo sido a escolha destes efetuada tendo em conta o fenótipo agressivo apresentado e a existência de familiares hipercolesterolémicos. Foi também tido em atenção na escolha a existência de familiares hipercolesterolémicos para todos os CI de forma a ser possível a realização de estudos de co-segregação das alterações nas famílias. Uma vez que a técnica de *exome sequencing* é ainda dispendiosa, esta foi apenas realizada a 5 CI. O envio destas amostras funcionou como um estudo preliminar, sendo o passo seguinte o envio de familiares destes CI de modo a se realizar uma análise por trio que poderá ser mais conclusiva. Um dos grandes desafios do *exome sequencing* foi a análise dos resultados, uma vez que para cada amostra foram obtidas cerca de 18 000 alterações. A primeira análise realizada foi a verificação da existência de alterações num dos genes associados à FH, tendo-se verificado que um dos CI apresentava uma alteração no gene *LDLR*. A pesquisa de alterações nos genes associados à FH foi realizada em todas as amostras enviadas - no entanto, uma vez que

para o estudo de *exome sequencing* o DNA deveria apresentar boa qualidade foi colhida nova amostra de sangue aos 5 CI e enviada a amostra da nova colheita. Após a análise dos dados e a verificação de uma alteração no gene *LDLR* num dos CI, foi realizada por sequenciação de Sanger a pesquisa da alteração em ambas as amostras de DNA desse CI, tendo-se verificado a presença da alteração na amostra enviada para sequenciação de exoma, mas não na amostra estudada anteriormente no âmbito do EPHF. Usualmente durante as colheitas e/ou extração de DNA é pedido ao profissional de saúde que tenha a máxima atenção para que não ocorra uma troca.

A sequenciação de exoma originou uma quantidade elevada de dados por amostra, facto que nos levou a optar por analisar apenas as alterações únicas nas 4 amostras. Contudo, continuou a verificar-se a existência de muitas alterações, pelo que seguidamente se optou por pesquisar apenas as alterações em genes descritos na literatura como estando associados ao metabolismo lipídico. Através de estudos de associação (GWAS, *Genome Wide Association Studies*) foi possível identificar novas variantes de DNA que influenciam os níveis de LDL no plasma. Uma das associações mais consistentes foi observada numa região do cromossoma 1p13, contendo os genes *SORT1*, *PSRC1*, *CELSR2*. No cromossoma 1, estão localizados também os genes *PCSK9*, *LDLRAP1*, reforçando a ideia que indivíduos com alterações nestes genes podem apresentar níveis de cLDL plasmático superiores aos indivíduos normolipidémicos, como foi descrito para o *SORT1*, embora não sejam tão elevados como os considerados para diagnóstico clínico de FH (cLDL >190 mg/dL nos adultos) (Kjolby *et al.* 2010; Linsel-Nitschke *et al.* 2010; Musunuru *et al.* 2010). Contudo os GWAS revelaram existir também associações com o valor de cLDL em regiões dos cromossomas 2, 8, 11, 18 e 19 (Holdt and Teupser 2013), onde se situam os genes *ABCG8*, *CPT1A*, *OSBPL1A* e *LDLR*.

Recentemente Talmud *et al.* (Talmud *et al.* 2013) publicaram um estudo onde desenvolveram um *score* baseado em SNPs associados a níveis elevados de cLDL, onde verificaram se a presença ou ausência deste SNPs tinham um efeito nos valores de cLDL. O objetivo era a aplicação deste *score* a indivíduos com diagnóstico clínico de FH mas sem uma alteração detetada de forma a explicar o fenótipo apresentado pelos indivíduos. Os SNPs considerados neste trabalho encontram-se distribuídos pelos genes: *PCSK9*; *CELSR2*; *APOB*; *ABCG8*; *SLC22A1*; *HFE*; *MYLIP*; *ST3GAL4*; *NYNRIN*; *LDLR* e *APOB*, dado que todos estes genes foram também contemplados na análise à sequenciação do exoma realizados no âmbito deste trabalho de doutoramento.



Através do programa de interação de proteínas, CYTOSCAPE, foram selecionadas as proteínas VEGFR, LRP2, LRPAP1, PDIA3 e DAB2 que interagem com a proteína LDLR. Todas as alterações encontradas nestes genes foram pesquisadas nos CI e seus familiares, tendo-se verificado que não existia co-segregação da alteração na família, para a maior parte das alterações.

Selecionaram-se 49 genes selecionados, tendo sido encontradas 50 alterações em 22 destes genes. O estudo molecular foi realizado para os CI e respectivos familiares a apenas 20 alterações distribuídas por 17 genes, tendo-se optado por não se efetuar o estudo das alterações sinónimas. Verificou-se que das 20 alterações nenhuma apresentava uma co-segregação clara que justificasse o fenótipo apresentado pelos CI. Como referido anteriormente o CI 28120 apresentava a alteração p.Asp19His no gene *ABCG8* justificando o fenótipo de litíases biliares apresentado pela avó materna, que tinha também a alteração, tendo sido a única alteração onde se encontrou uma correspondência clara entre o genótipo e fenótipo. No entanto algumas alterações nos genes *FLT1* e *SORT1* poderão estar também associadas a um fenótipo de hipercolesterolemia. Serão discutidos apenas os casos que foram considerados mais interessantes.

O gene *SORT1* codifica a sortilina, uma proteína multifuncional cuja relevância biológica tem vindo a ser cada vez mais valorizada, sendo conhecidas as suas funções como recetor de vários ligandos. Recentemente começaram a aparecer indícios de que a sortilina pode influenciar os níveis plasmáticos de cLDL e, conseqüentemente, o risco de DCV (Kathiresan *et al.* 2008; Coutinho *et al.* 2013). Em 2010 foram publicados quase em simultâneo três estudos desenvolvidos por grupos independentes (Kjolby *et al.* 2010; Linsel-Nitschke *et al.* 2010; Musunuru *et al.* 2010) que procuravam explorar o mecanismo biológico que poderia relacionar a sortilina com os níveis de cLDL. Para tal, utilizaram abordagens mecanisticamente diferentes tendo chegado a resultados paradoxais. Enquanto estudos independentes realizados por Linsel-Nitschke *et al.* e Musunuru *et al.* verificaram que a sobre-expressão *SORT1* originava um aumento da internalização de partículas cLDL, com conseqüente diminuição dos respetivos níveis plasmáticos (Linsel-Nitschke *et al.* 2010; Musunuru *et al.* 2010), os resultados obtidos por Kjolby *et al.* (Kjolby *et al.* 2010) verificaram que a deficiência de sortilina levava a uma redução de 50% na secreção de lipoproteínas e que a sua sobre-expressão se traduzia num aumento de 50% na secreção das mesmas proteínas, ou seja, os resultados

apontavam para uma correlação positiva entre a expressão do gene *SORT1* e a concentração de cLDL. A alteração p.Glu444Gln no gene *SORT1* foi encontrada no CI 25043, tendo-se verificado que a co-segregação da alteração na família não se verifica na prima hipercolesterolemica. Os resultados obtidos para esta alteração nesta família não são conclusivos, mas tendo em conta um artigo publicado recentemente por Tveten *et al.*, alterações no gene *SORT1* podem influenciar os níveis de cLDL plasmático, mas não são a causa de hipercolesterolemia familiar (Tveten *et al.* 2012). A função do gene *CELSR2* não é ainda conhecida, não se sabendo por isso qual o papel do *CELSR2* no metabolismo lipídico, nomeadamente na via do recetor das LDL. O estudo molecular das alterações p.Gln262Glu e p.Ala2689Ser realizado aos CI e seus familiares verificou não existir co-segregação com a hipercolesterolemia, não devendo ser esta a causa genética nestas 2 famílias. Apesar de não se localizar na mesma região cromossómica que os genes *SORT1*, *CELSR2* e *PSRC1*, o gene *SOAT1* encontra-se, também, no cromossoma 1. Este gene, também conhecido CoA:colesterol aciltransferase 1 (ACAT1), converte o colesterol sintetizado no fígado em ésteres de colesterol (Pramfalk *et al.* 2007). Neste estudo foi identificada a alteração p.Gln526Arg no gene *SOAT1*, em heterozigotia no CI 10146, uma vez que o pai hipercolesterolémico era homozigótico para esta alteração. Contudo, esta alteração foi detetada com uma percentagem de 12% no 1000 Genomes e 5,4% no ESP (Abecasis *et al.* 2012), em estudos de sequenciação de exoma, não devendo ser esta razão a causa da hipercolesterolemia no pai do CI 10146.

Nos CI 10146 e 25043 verificou-se a existência de alterações no gene *FLT1*. Após o estudo familiar verificou-se existir co-segregação com a hipercolesterolemia para a p.Arg1045Gln na família do CI 10146. Esta alteração encontra-se reportada na base de dados do *Exome Sequencing Project* (<http://evs.gs.washington.edu/EVS/>), tendo sido detetada apenas num indivíduo e não se conhecendo os seus valores bioquímicos. Na família do CI 25043 não se verifica co-segregação da alteração na tia hipercolesterolemica - no entanto esta alteração não foi descrita anteriormente. O gene *FLT1* codifica para a proteína recetor do fator de crescimento endotelial vascular (*vascular endothelial growth factor receptor* -VEGFR), desempenhando um papel importante na angiogénese e vasculogénese, mas o papel desta proteína na aterosclerose não é claro (Usui *et al.* 2007). A expressão deste recetor é detetada em células endoteliais vasculares, células da placenta e nos monócitos do sangue periférico. Um

estudo realizado por Usui *et al.* especulou que a proteína VEGFR fosse estimulada por partículas de LDL para a migração dos macrófagos, estando diretamente relacionada com o processo de aterosclerose na parede celular (Usui *et al.* 2007). Estudos posteriores verificaram que o efeito do gene *FLT1* no processo de desenvolvimento aterosclerótico será mínimo (Sandhofer *et al.* 2009).

No CI 28120, era já conhecida a mutação p.Gln4494del no gene *APOB* quando foi realizado o estudo de sequenciação do exoma - no entanto não tinha sido realizado os estudos funcionais, motivo pelo qual foi selecionado para este estudo. Verificou-se que este CI apresentava também as alterações p.Val667Met, p.Asp19His e p.Arg296Cys nos genes *LRP5*, *ABCG8* e *DAB2*, respetivamente. A alteração p.Asp19His foi descrita como estando associada a um aumento da suscetibilidade de desenvolver litíases biliares de colesterol, tendo-se verificado que a avó do CI que apresentava também a alteração já tinha realizado uma exérese da vesícula, como discutido anteriormente. Neste CI verificou-se também a presença da alteração p.Val667Met do gene *LRP5* que não co-segrega com a hipercolesterolemia na família. Neste CI foi também encontrada a alteração p.Arg296Cys no gene *DAB2* que está presente nos indivíduos com fenótipo de hipercolesterolemia. O gene *DAB2*, codifica para uma proteína que apresenta uma organização estrutural e de domínios muito semelhante à *LDLRAP1* sabendo-se que em certos tipos de células pode ser esta a proteína adaptadora que permite a formação de vesículas revestidas de clatrina e ocorra a endocitose.

A abordagem utilizada neste trabalho não contemplou a análise de exoma a CI e familiares que poderia ajudar a encontrar uma alteração que explicasse o fenótipo de hipercolesterolemia - em estudos futuros esta abordagem deverá ser considerada. Uma vez que o *exome sequencing* origina uma elevada quantidade dados, a análise de todos os resultados obtidos continuará posteriormente a este trabalho, na tentativa da identificação da causa genética da FH nestes indivíduos, incluindo o estudo dos exomas dos familiares de modo a realizarem novos filtros que podem levar a novas variantes em genes não associados ao metabolismo lipídico.

Quando foi realizado o estudo de sequenciação do exoma aos 5 CI do EPHF, no decorrer deste trabalho de doutoramento, nenhum estudo de sequenciação de exoma ou genoma tinha sido descrito na literatura em indivíduos com diagnóstico clínico de FH sem mutação num dos genes associados à FH. No ano de 2012, surgiram os primeiros trabalhos sobre sequenciação de exoma na FH (Futema *et al.* 2012; Motazacker *et al.*

2012). O primeiro estudo foi realizado num grupo holandês numa família sem mutação nos genes associados à FH. No entanto, este grupo realizou em paralelo uma análise de *linkage* tendo cruzado os dados obtidos em ambas as análises e verificado que o número de alterações diminuía consideravelmente. A única alteração encontrada pelo grupo holandês foi uma mutação no gene *APOB* localizada numa região que não é usualmente estudada. Também em 2012, um grupo inglês realizou um estudo de *targed exome sequencing* em 48 CI com diagnóstico clínico de FH e sem mutação, tendo-se verificado uma alteração num dos 3 genes associados à FH em 20 destes CI. Com o aparecimento das NGS, sendo o seu custo cada vez mais acessível, estes métodos de rastreio a partir de *chips* com mutações específicas foram descontinuados, estando a maioria dos grupos a desenvolver técnicas para o diagnóstico por *targed exome sequencing*. O facto de algumas *cohorts* realizarem o diagnóstico por métodos onde a análise ao *LDLR* não era realizada na totalidade pode estar a falsear o número de CI identificados nos países onde este tipo de diagnóstico era aplicado. Recentemente, Thomas *et al* (Thomas *et al.* 2013) publicaram um estudo onde descrevem a análise realizada a um CI e seus familiares através de sequenciação de exoma em conjunto com estudos de *linkage* na família, semelhante ao publicado anteriormente pelo grupo holandês. Curiosamente também este grupo verificou que a causa genética da hipercolesterolemia na família era uma mutação no gene *APOB* no exão 3, fora da região de ligação ao ligando, não tendo sido esta alteração comprovada funcionalmente (Thomas *et al.* 2013).

Alguns dos grupos que estudam a FH em todo o mundo, realizaram estudos por *exome sequencing* em indivíduos com diagnóstico clínico de FH e sem mutação, tendo verificado apenas a existência de alterações em genes já descritos e associados à FH, levando a pensar que não haverá outro gene associado à FH, mas sim mutações em locais não estudados habitualmente, como é o caso do gene *APOB*. A verificar-se a existência de outro gene, irá estar associado a uma pequena percentagem de casos. Não pode ser descartada a hipótese de a maioria dos doentes apresentarem uma dislipidémia poligénica.

Através do estudo *exome sequencing* nos 5 CI identificou-se 2 alterações no gene *FLT1* em 2 CI não relacionados e 1 alteração no gene *SORT1* noutra CI que co-segregam com a hipercolesterolemia na família. Os estudos funcionais das alterações encontradas nestes genes serão importante para classificar as mesmas quanto à sua patogenicidade. No caso das alterações encontradas nestes genes serem as responsáveis pelo fenótipo de

hipercolesterolemia nestes indivíduos serão as primeiras alterações descritas originando FH depois da descoberta do PCSK9 à 10 anos atrás. Para além destas 3 alterações, a sequenciação do exoma permitiu identificar uma alteração no gene *ABCG8* no CI 28120 e em 2 familiares, clarificando o fenótipo apresentado na família. Até ao momento os resultados obtidos neste trabalho são os primeiros que apontam para que possa existir alterações noutros associados à FH.

Apesar de cerca de 50% dos indivíduos com diagnóstico clínico de FH não apresentarem uma alteração que justifique o seu fenótipo (neste trabalho assim como se encontra descrito noutras populações), a causa da hipercolesterolemia nestes indivíduos continua por justificar. As novas tecnologias de sequenciação vieram trazer uma nova abordagem metodológica importante para a identificação de novas alterações em genes e/ou regiões exónicas que não eram habitualmente estudadas; no entanto é ainda necessário desenvolver novas abordagens de análise que ajudem a esclarecer a causa da doença em indivíduos com diagnóstico clínico de FH. O aumento de casos com FH identificados é importante uma vez que estes apresentam um risco elevado de sofrerem uma DCV prematura. À semelhança do diagnóstico de outras patologias (Daiger *et al.* 2014; Yoshimura *et al.* 2014) o diagnóstico da FH também irá ser reformulado de modo a potenciar as vantagens das novas tecnologias de sequenciação, dado que *target sequencing* de genes associados ao metabolismo lipídico será a metodologia de escolha ou a mesmo a sequenciação do exoma caso o custo reduza substancialmente.

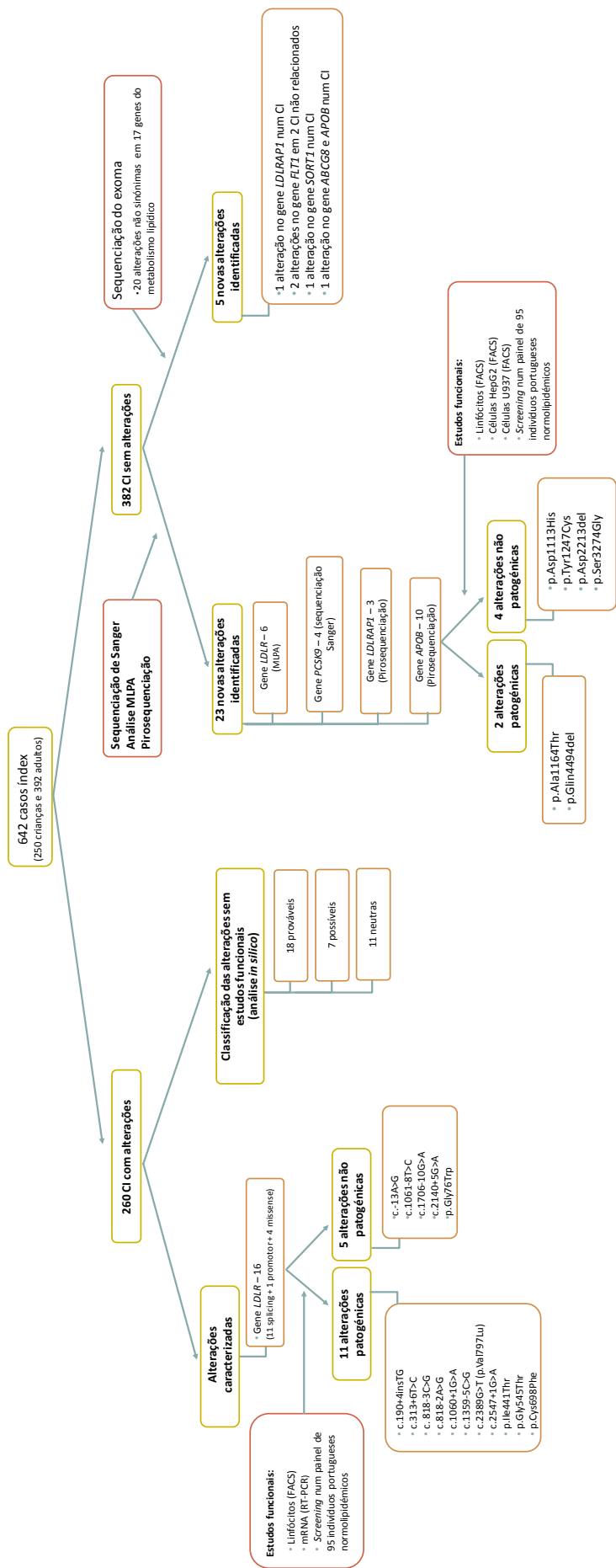


Figura IV.1 – Esquemaização dos resultados moleculares mais relevantes.







## CAPÍTULO V

---

# Conclusões e Perspetivas Futuras



## V CONCLUSÕES E PERSPETIVAS FUTURAS

A Hipercolesterolemia Familiar foi a primeira doença genética do metabolismo lipídico a ser caracterizada molecularmente (Brown and Goldstein 1986). A maioria dos doentes com FH morre precocemente devido a doença coronária que poderia ser evitada ou minimizada se o doente fosse corretamente diagnosticado de preferência em idade jovem, o que só acontece quando existe um correto diagnóstico molecular. Estudos realizados demonstram que jovens com FH, precocemente identificados e tratados, podem nunca vir a desenvolver qualquer tipo de doença coronária (Hopkins 2010). Geneticamente esta patologia caracteriza-se por mutações em três genes: *LDLR*, *APOB* e *PCSK9*. No entanto, cerca de 65% dos doentes com diagnóstico clínico de FH não apresentam nenhuma mutação funcional num destes genes (Talmud *et al.* 2013).

De forma a identificar e clarificar a causa genética de indivíduos com diagnóstico clínico de FH o primeiro objetivo deste trabalho foi a reorganização do estudo molecular no EPHF. No decorrer desta reorganização foram testadas novas abordagens metodológicas, nomeadamente a pirosequenciação, uma técnica de sequenciação de nova geração, de forma a agilizar e otimizar o diagnóstico molecular da FH para que este se tornasse menos moroso e dispendioso. A primeira abordagem metodológica, realizada em 2009 através do equipamento *454 Life Science*, teve como objetivo resequenciar o gene *LDLR* de 65 doentes com diagnóstico clínico de FH sem mutação identificada, bem como a sequenciação de todo o gene *APOB*, que não era prática comum em laboratórios de rotina. Com esta abordagem, identificaram-se 10 possíveis mutações no gene *APOB*, todas situadas fora da região postulada *consensus* de ligação do ligando ao LDLR. Em 2012, foi realizada uma segunda abordagem metodológica, através do equipamento *GS Junior system*, tendo-se verificado uma melhoria significativa nos resultados obtidos. A possibilidade de realizar o diagnóstico clínico através de *target sequencing* ficou evidenciada com a abordagem pelo equipamento *GS Junior system*. Neste momento, no EPHF encontra-se em fase de implementação o diagnóstico genético da FH através de *target sequencing* pelo equipamento *MiSEQ*.

O estudo funcional de alterações encontradas nos genes *APOB* e *LDLR* era outro dos objetivos propostos. O método escolhido para os ensaios funcionais em ambos os genes necessitava de uma nova colheita de sangue aos indivíduos que apresentavam as alterações, tendo sido um fator limitante para o número de alterações caracterizadas

funcionalmente. Para o gene *LDLR* foi realizado o estudo funcional a 12 alterações de splicing e 4 pontuais. Os estudos funcionais nas alterações de *splicing* foram realizados a partir do mRNA de células mononucleares dos indivíduos e as alterações pontuais a partir dos linfócitos dos indivíduos que apresentavam as alterações. Verificou-se que 12 eram patogénicas e que 4 não originavam doença. Para as alterações de *splicing*, apesar de se ter realizado uma análise funcional, seria importante em trabalhos futuros estudar as alterações por *real time* PCR de modo a quantificar os transcritos obtidos para cada alelo, para se conhecer o real efeito de cada alteração em cada indivíduo. Os estudos funcionais a 4 alterações no gene *APOB* foram realizados a partir do isolamento das partículas de LDL dos indivíduos com as alterações. Comprovou-se a patogenicidade das primeiras mutações fora da região postulada *consensus* de ligação do ligando ao LDLR, sendo ainda necessário realizar estudos funcionais a 4 alterações. A realização destes estudos está prevista para o ano 2014. Usualmente, o diagnóstico molecular da FH era realizado apenas a uma pequena parte dos exões 26 e 29 do gene *APOB*, mas com a comprovação da patogenicidade das mutações descritas neste trabalho, fica demonstrada a necessidade de se alargar o estudo a todo o gene *APOB*, dado que as novas metodologia de sequenciação em larga escala facilitarão este estudo molecular.

Encontram-se descritas mais de 1500 alterações no gene *LDLR* em todo o mundo, contudo cerca de 40% não apresentam estudos funcionais que comprovem, ou não, a sua patogenicidade (Usifo *et al.* 2012). Na impossibilidade da realização de estudos funcionais a todas as alterações descritas, foi proposta neste trabalho de doutoramento uma classificação de modo a facilitar a catalogação das alterações. Para tal foi desenvolvido um critério com base em 6 programas bioinformáticos para alterações pontuais e 3 programas para alterações de *splicing* que preveem o efeito que as alterações originam na proteína, critério este que foi validado por comparação com uma análise *in vitro* de algumas mutações anteriormente caracterizadas. Obteve-se uma sensibilidade e uma especificidade de 86% e 100% respetivamente para as alterações de *missense* e de 100% e 100% para alterações de *splicing*, demonstrando a eficiência da classificação desenvolvida neste trabalho.

Com base nesta classificação, todas as alterações sem estudos funcionais foram catalogadas como provavelmente patogénicas, possivelmente patogénicas ou neutras. Também os indivíduos estudados no EPHF foram classificados quanto ao seu diagnóstico molecular com base nas alterações que apresentavam. Esta classificação

permitiu que fosse feita uma divisão quanto ao tipo de diagnóstico: definitivo, provável, possível e neutro.

Seguindo esta classificação, todos os indivíduos homozigóticos e heterozigóticos compostos identificados no EPHF com alterações no gene *LDLR* (3 indivíduos homozigóticos e 5 heterozigóticos compostos) foram analisados. Após a classificação das alterações sem estudos funcionais, constatou-se que 2 indivíduos eram de facto homozigóticos e que apenas um indivíduo era heterozigótico composto, sendo os restantes classificados como tendo um diagnóstico possível de FH homozigótica.

Um dos objetivos propostos neste trabalho era a realização de uma análise estatística na qual se pretendia comparar os parâmetros bioquímicos entre indivíduos com diagnóstico clínico de FH com e sem mutação. A análise foi realizada em dois grupos: pediátrico e adulto. Dentro de cada grupo existia ainda uma divisão consoante o diagnóstico molecular, o que permitiu verificar que os indivíduos com alterações classificadas como provavelmente patogénicas apresentam valores bioquímicos mais semelhantes aos indivíduos com mutações patogénicas e estatisticamente diferentes dos indivíduos com alterações possivelmente patogénicas, neutras ou dos indivíduos sem alterações. Neste trabalho, foi realizada para todos os indivíduos referenciados ao EPHF uma caracterização a 7 parâmetros lipídicos (CT, cLDL, cHDL, TG, apoA-I, apoB, Lp(a)), não sendo esta prática usual nos restantes laboratórios que estudam a FH. Para além da análise destes parâmetros, foram também analisados mais 5 parâmetros (apoA-II, apoC-II, apoC-III, apoE e sdLDL), de forma a se obter uma melhor caracterização bioquímica dos indivíduos. Verificou-se que a caracterização bioquímica complementar é importante e informativa, podendo ser fundamental para o diagnóstico clínico dos indivíduos referenciados ao estudo da FH em todo o mundo. Uma vez que as crianças parecem apresentar um fenótipo menos influenciado por causas externas (erros alimentares, tabagismo, sedentarismo), foi pesquisado um biomarcador que permitisse distinguir uma hipercolesterolemia monogénica de uma poligénica/externa. Verificou-se que a razão apoB/apoA1, com um valor de corte  $\geq 0,68$ , é o melhor biomarcador para diferenciar a FH de outras dislipidemias. O estudo para a identificação de um biomarcador eficaz na diferenciação das hipercolesterolemias foi apenas realizado em crianças, mas pretende-se em trabalhos futuros estender esta abordagem ao grupo adulto.

O último objetivo deste trabalho de doutoramento era a realização de sequenciação de exoma a 5 indivíduos com diagnóstico clínico de FH mas sem uma mutação identificada

num dos 3 genes associados à FH. A sequenciação de exoma originou uma quantidade elevada de dados por amostra, tendo sido necessária a aplicação de várias estratégias de modo a obter um número aceitável de variantes a analisar. Uma das estratégias foi a seleção de 49 genes associados ao metabolismo lipídico, verificando-se a existência de 50 alterações em 22 destes genes. O estudo molecular foi realizado para os CI e respetivos familiares a apenas 20 alterações distribuídas por 17 genes uma vez que se optou por não se alargar o estudo às alterações sinónimas. Verificou-se que das 20 alterações, apenas um CI apresentava a alteração p.Asp19His no gene *ABCG8* justificando o fenótipo de litíases biliares apresentado pela avó materna, que apresentava também a alteração, tendo-se encontrado uma correspondência clara entre o genótipo e fenótipo. A identificação desta alteração permitiu também compreender melhor o fenótipo da família com a mutação no gene *APOB*. Neste estudo identificaram-se também 2 alterações no gene *FLT1* em 2 CI não relacionados e 1 alteração no gene *SORT1* noutro CI que co-segregavam com a hipercolesterolemia na família. Os estudos funcionais das alterações nos genes *SORT1* e *FLT1* serão importantes para classificar as mesmas quanto à sua patogenicidade. Caso se comprove que uma destas alterações é patogénica estaremos perante novos genes associados à FH. Nos últimos anos e apesar da sequenciação do exoma e genoma realizada por grupos de pesquisa especializados neste tema, bem como o estudo realizado neste trabalho, não foi possível identificar novos genes associados à FH. Desta forma em trabalhos futuros, deveria ser estudada a interação de microRNAs (miR) na homeostase lipídica para se verificar qual o envolvimento destes no metabolismo lipídico e se podem perturbar o metabolismo do colesterol em doentes com FH.

No decorrer deste trabalho doutoral, foi possível caracterizar alterações no gene *LDLR* cuja patogenicidade não era conhecida; identificaram-se e caracterizaram-se as primeiras mutações fora da região postulada *consensus* de ligação do ligando ao LDLR bem como mutações no gene *APOB* em locais fora da região *consensus*. Na impossibilidade de caracterizar todas as alterações encontradas no gene *LDLR*, foi elaborado um método de classificação que permitiu distinguir as alterações neutras das alterações prováveis, levando a uma melhor caracterização bioquímica dos indivíduos referenciados ao EPHF com diagnóstico clínico de FH. Através do estudo do exoma foram identificadas possíveis alterações patogénicas em genes associados ao metabolismo lipídico, mas não diretamente à FH, *FLT1* e *SORT1*, podendo ser indicativo de um novo gene associado a esta doença. Através dos resultados obtidos no

decorrer deste trabalho foi possível contribuir para um aumento científico no que diz respeito à base genética da FH.





## REFERÊNCIAS

---



## REFERÊNCIAS

- Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA. 2012. An integrated map of genetic variation from 1,092 human genomes. *Nature* 491: 56–65.
- Abifadel M, Rabès J-P, Devillers M, Munnich A, Erlich D, Junien C, Varret M, Boileau C. 2009. Mutations and polymorphisms in the proprotein convertase subtilisin kexin 9 (PCSK9) gene in cholesterol metabolism and disease. *Hum. Mutat.* 30: 520–9.
- Abifadel M, Varret M, Rabes JP, Allard D, Ouguerram K, Devillers M, Cauaud C, Benjannet S, Wickham L, Erlich D, Derré A, Villéger L, et al. 2003. Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nat Genet* 34: 154–156.
- Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. 2010. A method and server for predicting damaging missense mutations. *Nat. Methods* 7: 248–9.
- Alves AC, Etxebarria A, Soutar AK, Martin C, Bourbon M. 2013. Novel functional APOB mutations outside LDL-binding region causing familial hypercholesterolemia. *Hum. Mol. Genet.*
- Alves AC, Medeiros AM, Francisco V, Gaspar IM, Rato Q, Bourbon M. 2010. Molecular diagnosis of familial hypercholesterolemia: an important tool for cardiovascular risk stratification. *Rev Port Cardiol* 29: 907–921.
- Arca M, Zuliani G, Wilund K, Campagna F, Fellin R, Bertolini S, Calandra S, Ricci G, Glorioso N, Maioli M, Pintus P, Carru C, et al. 2002. Autosomal recessive hypercholesterolaemia in Sardinia, Italy, and mutations in ARH: a clinical and molecular genetic analysis. *Lancet* 359: 841–7.
- Badano JL, Katsanis N. 2002. Beyond Mendel: an evolving view of human genetic disease transmission. *Nat. Rev. Genet.* 3: 779–89.
- Bailey J. 1973. *Atherogenesis: Initiating Factors*. Amsterdam: Elsevier, Amsterdam: Ciba Foundation Symposium. 402–407.
- Barson JR, Karatayev O, Chang G-Q, Johnson DF, Bocarsly ME, Hoebel BG, Leibowitz SF. 2009. Positive relationship between dietary fat, ethanol intake, triglycerides, and hypothalamic peptides: counteraction by lipid-lowering drugs. *Alcohol* 43: 433–41.
- Bednarska-Makaruk M, Broda G, Kurjata P, Rodo M, Roszczyńko M, Rywik S, Wehr H. 2001. Apolipoprotein E genotype, lipid levels and coronary heart disease in a Polish population group. *Eur. J. Epidemiol.* 17: 789–92.
- Benjannet S, Rhainds D, Essalmani R, Mayne J, Wickham L, Jin W, Asselin M-C, Hamelin J, Varret M, Allard D, Trillard M, Abifadel M, et al. 2004. NARC-1/PCSK9 and its natural mutants: zymogen cleavage and effects on the low density lipoprotein (LDL) receptor and LDL cholesterol. *J. Biol. Chem.* 279: 48865–75.

- Benn M, Watts GF, Tybjaerg-Hansen A, Nordestgaard BG. 2012. Familial hypercholesterolemia in the danish general population: prevalence, coronary artery disease, and cholesterol-lowering medication. *J. Clin. Endocrinol. Metab.* 97: 3956–64.
- Berge KE, Ose L, Leren TP. 2006. Missense mutations in the PCSK9 gene are associated with hypocholesterolemia and possibly increased response to statin therapy. *Arterioscler. Thromb. Vasc. Biol.* 26: 1094–100.
- Bertolini S, Cassanelli S, Garuti R, Ghisellini M, Simone ML, Rolleri M, Masturzo P, Calandra S. 1999. Analysis of LDL receptor gene mutations in Italian patients with homozygous familial hypercholesterolemia. *Arterioscler. Thromb. Vasc. Biol.* 19: 408–18.
- Bertolini S, Garuti R, Lelli W, Rolleri M, Tiozzo RM, Ghisellini M, Simone ML, Masturzo P, Elicio NC, Stefanutti C. 1995. Four novel partial deletions of LDL-receptor gene in Italian patients with familial hypercholesterolemia. *Arterioscler. Thromb. Vasc. Biol.* 15: 81–8.
- Boren J, Ekstrom U, Agren B, Nilsson-Ehle P, Innerarity TL. 2001. The molecular mechanism for the genetic disorder familial defective apolipoprotein B100. *J Biol Chem* 276: 9214–9218.
- Bourbon M, Alves AC, Medeiros AM, Silva S, Soutar AK. 2008. Familial hypercholesterolaemia in Portugal. *Atherosclerosis* 196: 633–642.
- Bourbon M, Duarte MA, Alves AC, Medeiros AM, Marques L, Soutar AK. 2009. Genetic diagnosis of familial hypercholesterolaemia: the importance of functional analysis of potential splice-site mutations. *J Med Genet* 46: 352–357.
- Bourbon M, Rato Q. 2006. Portuguese Familial Hypercholesterolemia Study: presentation of the study and preliminary results. *Rev Port Cardiol* 25: 999–1013.
- Bourbon M, Sun XM, Soutar AK. 2007. A rare polymorphism in the low density lipoprotein (LDL) gene that affects mRNA splicing. *Atherosclerosis* 195: e17–20.
- Broek AJ Van den, Hollaar L, Schaefer HI, Laarse A Van der, Schuster H, Defesche JC, Kastelein JJ, 't Hooft FM Van. 1994. Screening for familial defective apolipoprotein B-100 with improved U937 monocytic proliferation assay. *Clin. Chem.* 40: 395–9.
- Brouwers MCGJ, Greevenbroek MMJ van, Stehouwer CDA, Graaf J de, Stalenhoef AFH. 2012. The genetics of familial combined hyperlipidaemia. *Nat. Rev. Endocrinol.* 8: 352–62.
- Brown MS, Goldstein JL. 1979. Receptor-mediated endocytosis: insights from the lipoprotein receptor system. *Proc Natl Acad Sci U S A* 76: 3330–3337.
- Brown MS, Goldstein JL. 1986. A receptor-mediated pathway for cholesterol homeostasis. *Science* (80-. ). 232: 34–47.
- Buch S, Schafmayer C, Völzke H, Becker C, Franke A, Eller-Eberstein H von, Kluck C, Bässmann I, Brosch M, Lammert F, Miquel JF, Nervi F, et al. 2007. A genome-wide association scan identifies the hepatic cholesterol transporter ABCG8 as a susceptibility factor for human gallstone disease. *Nat. Genet.* 39: 995–9.
- Burton PR, Tobin MD, Hopper JL. 2005. Key concepts in genetic epidemiology. *Lancet* 366: 941–51.

- Cameron J, Holla ØL, Ranheim T, Kulseth MA, Berge KE, Leren TP. 2006. Effect of mutations in the PCSK9 gene on the cell surface LDL receptors. *Hum. Mol. Genet.* 15: 1551–8.
- Catapano AL, Reiner Z, Backer G De, Graham I, Taskinen MR, Wiklund O, Agewall S, Alegria E, Chapman M, Durrington P, Erdine S, Halcox J, et al. 2011a. ESC/EAS Guidelines for the management of dyslipidaemias The Task Force for the management of dyslipidaemias of the European Society of Cardiology (ESC) and the European Atherosclerosis Society (EAS). *Atherosclerosis* 217: 3–46.
- Catapano AL, Reiner Z, Backer G De, Graham I, Taskinen MR, Wiklund O, Agewall S, Alegria E, Chapman MJ, Durrington P, Erdine S, Halcox J, et al. 2011b. ESC/EAS Guidelines for the management of dyslipidaemias: the Task Force for the management of dyslipidaemias of the European Society of Cardiology (ESC) and the European Atherosclerosis Society (EAS). *Atherosclerosis* 217 Suppl : S1–44.
- Chang J-H, Pan J-P, Tai D-Y, Huang A-C, Li P-H, Ho H-L, Hsieh H-L, Chou S-C, Lin W-L, Lo E, Chang C-Y, Tseng J, et al. 2003. Identification and characterization of LDL receptor gene mutations in hyperlipidemic Chinese. *J. Lipid Res.* 44: 1850–8.
- Chatterton JE, Schlapfer P, Büttler E, Gutierrez MM, Puppione DL, Pullinger CR, Kane JP, Curtiss LK, Schumaker VN. 1995. Identification of apolipoprotein B100 polymorphisms that affect low-density lipoprotein metabolism: description of a new approach involving monoclonal antibodies and dynamic light scattering. *Biochemistry* 34: 9571–80.
- Chen JJ, Berlin FS, Margolis S. 1986. Effect of large-dose progesterone on plasma levels of lipids, lipoproteins and apolipoproteins in males. *J Endocrinol Invest* 9: 281–285.
- Chen Z-C, Shin S-J, Kuo K-K, Lin K-D, Yu M-L, Hsiao P-J. 2008. Significant association of ABCG8:D19H gene polymorphism with hypercholesterolemia and insulin resistance. *J. Hum. Genet.* 53: 757–63.
- Chiou K-R, Charng M-J. 2010. Detection of mutations and large rearrangements of the low-density lipoprotein receptor gene in Taiwanese patients with familial hypercholesterolemia. *Am. J. Cardiol.* 105: 1752–8.
- Cohen J, Pertsemlidis A, Kotowski IK, Graham R, Garcia CK, Hobbs HH. 2005. Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. *Nat Genet* 37: 161–165.
- Cohen JC, Boerwinkle E, Mosley Jr. TH, Hobbs HH. 2006. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. *N Engl J Med* 354: 1264–1272.
- Collins FS, Morgan M, Patrinos A. 2003. The Human Genome Project: lessons from large-scale biology. *Science* 300: 286–90.
- Coon H, Eckfeldt JH, Leppert MF, Myers RH, Arnett DK, Heiss G, Province MA, Hunt SC. 2002. A genome-wide screen reveals evidence for a locus on chromosome 11 influencing variation in LDL cholesterol in the NHLBI Family Heart Study. *Hum. Genet.* 111: 263–9.
- Coon H, Leppert MF, Eckfeldt JH, Oberman A, Myers RH, Peacock JM, Province MA, Hopkins PN, Heiss G. 2001. Genome-wide linkage analysis of lipids in the Hypertension Genetic Epidemiology Network (HyperGEN) Blood Pressure Study. *Arterioscler. Thromb. Vasc. Biol.* 21: 1969–76.

- Cotton RG, Scriver CR. 1998. Proof of “disease causing” mutation. *Hum. Mutat.* 12: 1–3.
- Coutinho MF, Bourbon M, Prata MJ, Alves S. 2013. Sortilin and the risk of cardiovascular disease. *Rev. Port. Cardiol.* 32: 793–799.
- Cunningham D, Danley DE, Geoghegan KF, Griffor MC, Hawkins JL, Subashi TA, Varghese AH, Ammirati MJ, Culp JS, Hoth LR, Mansour MN, McGrath KM, et al. 2007. Structural and biophysical studies of PCSK9 and its mutants linked to familial hypercholesterolemia. *Nat. Struct. Mol. Biol.* 14: 413–9.
- Daiger SP, Bowne SJ, Sullivan LS, Blanton SH, Weinstock GM, Koboldt DC, Fulton RS, Larsen D, Humphries P, Humphries MM, Pierce EA, Chen R, et al. 2014. Application of Next-Generation Sequencing to Identify Genes and Mutations Causing Autosomal Dominant Retinitis Pigmentosa (adRP). *Adv. Exp. Med. Biol.* 801: 123–9.
- Daniels SR, Gidding SS, Ferranti SD de. 2011. Pediatric aspects of familial hypercholesterolemias: recommendations from the National Lipid Association Expert Panel on Familial Hypercholesterolemia. *J. Clin. Lipidol.* 5: S30–7.
- Davis CG, Goldstein JL, Südhof TC, Anderson RG, Russell DW, Brown MS. 1987. Acid-dependent ligand dissociation and recycling of LDL receptor mediated by growth factor homology region. *Nature* 326: 760–5.
- Davis CG, Lehrman MA, Russell DW, Anderson RG, Brown MS, Goldstein JL. 1986. The J.D. mutation in familial hypercholesterolemia: amino acid substitution in cytoplasmic domain impedes internalization of LDL receptors. *Cell* 45: 15–24.
- Dawn Teare M, Barrett JH. 2005. Genetic linkage studies. *Lancet* 366: 1036–44.
- Day IN, Whittall RA, O’Dell SD, Haddad L, Bolla MK, Gudnason V, Humphries SE. 1997. Spectrum of LDL receptor gene mutations in heterozygous familial hypercholesterolemia. *Hum. Mutat.* 10: 116–27.
- Dean M, Allikmets R. 1995. Evolution of ATP-binding cassette transporter genes. *Curr. Opin. Genet. Dev.* 5: 779–85.
- Dedoussis GVZ, Schmidt H, Genschel J. 2004. LDL-receptor mutations in Europe. *Hum. Mutat.* 24: 443–59.
- Defesche JC, Schuurman EJM, Klaaijsen LN, Khoo KL, Wiegman A, Stalenhoef AFH. 2008. Silent exonic mutations in the low-density lipoprotein receptor gene that cause familial hypercholesterolemia by affecting mRNA splicing. *Clin. Genet.* 73: 573–8.
- DeMott K, Nherera L, Shaw EJ, Minhas R, Humphries SE, Kathoria M, Ritchie G, Nunes V, Davies D, Lee P, McDowell I, Neil A, et al. 2008. Clinical guidelines and evidence review for familial hypercholesterolaemia: the identification and management of adults and children with familial hypercholesterolaemia. . London: National Collaborating Centre for Primary Care and Royal College of General Practitioners.
- Dogan RI, Getoor L, Wilbur WJ, Mount SM. 2007. SplicePort--an interactive splice-site analysis tool. *Nucleic Acids Res.* 35: W285–91.

- Dvir H, Shah M, Girardi E, Guo L, Farquhar MG, Zajonc DM. 2012. Atomic structure of the autosomal recessive hypercholesterolemia phosphotyrosine-binding domain in complex with the LDL-receptor tail. *Proc. Natl. Acad. Sci. U. S. A.* 109: 6916–21.
- Eden ER, Sun X-M, Patel DD, Soutar AK. 2007. Adaptor protein disabled-2 modulates low density lipoprotein receptor synthesis in fibroblasts from patients with autosomal recessive hypercholesterolaemia. *Hum. Mol. Genet.* 16: 2751–9.
- Erlich Y, Mitra PP, delaBastide M, McCombie WR, Hannon GJ. 2008. Alta-Cyclic: a self-optimizing base caller for next-generation sequencing. *Nat. Methods* 5: 679–82.
- Esser V, Russell DW. 1988. Transport-deficient mutations in the low density lipoprotein receptor. Alterations in the cysteine-rich and cysteine-poor regions of the protein block intracellular transport. *J. Biol. Chem.* 263: 13276–81.
- Etxebarria A, Palacios L, Stef M, Tejedor D, Uribe KB, Oleaga A, Irigoyen L, Torres B, Ostolaza H, Martin C. 2012. Functional characterization of splicing and ligand-binding domain variants in the LDL receptor. *Hum Mutat* 33: 232–243.
- Fasano T, Sun X-M, Patel DD, Soutar AK. 2009. Degradation of LDLR protein mediated by “gain of function” PCSK9 mutants in normal and ARH cells. *Atherosclerosis* 203: 166–71.
- Fouchier SW, Defesche JC, Umans-Eckenhuis MW, Kastelein JP. 2001. The molecular basis of familial hypercholesterolemia in The Netherlands. *Hum. Genet.* 109: 602–15.
- Fouchier SW, Kastelein JJP, Defesche JC. 2005. Update of the molecular basis of familial hypercholesterolemia in The Netherlands. *Hum. Mutat.* 26: 550–6.
- Frazier ME, Johnson GM, Thomassen DG, Oliver CE, Patrinos A. 2003. Realizing the potential of the genome revolution: the genomes to life program. *Science* 300: 290–3.
- Fredrickson DS. 1974. Plasma lipoproteins and apolipoproteins. *Harvey Lect.* 68: 185–237.
- Fredrickson DS, Levy RI, Lees RS. 1967. Fat transport in lipoproteins--an integrated approach to mechanisms and disorders. *N Engl J Med* 276: 94–103 contd.
- Frostedgård J, Hamsten A, Gidlund M, Nilsson J. 1990. Low density lipoprotein-induced growth of U937 cells: a novel method to determine the receptor binding of low density lipoprotein. *J. Lipid Res.* 31: 37–44.
- Futema M, Plagnol V, Whittall RA, Neil HAW, Humphries SE. 2012. Use of targeted exome sequencing as a diagnostic tool for Familial Hypercholesterolaemia. *J. Med. Genet.* 49: 644–9.
- Garcia CK, Wilund K, Arca M, Zuliani G, Fellin R, Maioli M, Calandra S, Bertolini S, Cossu F, Grishin N, Barnes R, Cohen JC, et al. 2001. Autosomal recessive hypercholesterolemia caused by mutations in a putative LDL receptor adaptor protein. *Science* (80-. ). 292: 1394–1398.
- Garuti R, Lelli N, Barozzini M, Tiozzo R, Ghisellini M, Simone ML, Li Volti S, Garozzo R, Mollica F, Vergoni W, Bertolini S, Calandra S. 1996. Two novel partial deletions of LDL-receptor gene in Italian patients with familial hypercholesterolemia (FH Siracusa and FH Reggio Emilia). *Atherosclerosis* 121: 105–17.

- Gazi IF, Tsimihodimos V, Tselepis AD, Elisaf M, Mikhailidis DP. 2007. Clinical importance and therapeutic modulation of small dense low-density lipoprotein particles. *Expert Opin. Biol. Ther.* 7: 53–72.
- Gilissen C, Hoischen A, Brunner HG, Veltman JA. 2012. Disease gene identification strategies for exome sequencing. *Eur. J. Hum. Genet.* 20: 490–7.
- Gillian-Daniel DL, Bates PW, Tebon A, Attie AD. 2002. Endoplasmic reticulum localization of the low density lipoprotein receptor mediates presecretory degradation of apolipoprotein B. *Proc. Natl. Acad. Sci. U. S. A.* 99: 4337–42.
- GOFMAN JW, GLAZIER F, TAMPLIN A, STRISOWER B, LALLA O DE. 1954. Lipoproteins, coronary heart disease, and atherosclerosis. *Physiol. Rev.* 34: 589–607.
- GOFMAN JW, JONES HB, LYON TP, LINDGREN F, GRAHAM D, STRISOWER B, NICHOLS A. 1952. Atherosclerosis, lipoproteins, and coronary artery disease. *Wis. Med. J.* 51: 687–9.
- Goldmann R, Tichý L, Freiburger T, Zapletalová P, Letocha O, Soska V, Fajkus J, Fajkusová L. 2010. Genomic characterization of large rearrangements of the LDLR gene in Czech patients with familial hypercholesterolemia. *BMC Med. Genet.* 11: 115.
- Goldstein JL, Brown MS. 1974. Binding and degradation of low density lipoproteins by cultured human fibroblasts. Comparison of cells from a normal subject and from a patient with homozygous familial hypercholesterolemia. *J Biol Chem* 249: 5153–5162.
- Goldstein JL, Brown MS. 1979. The LDL receptor locus and the genetics of familial hypercholesterolemia. *Annu Rev Genet* 13: 259–289.
- Goldstein JL, Brown MS. 1985. Familial hypercholesterolemia: a genetic receptor disease. *Hosp Pr. (Off Ed)* 20: 35–41,45–46.
- Goldstein JL, Brown MS, Anderson RG, Russell DW, Schneider WJ. 1985. Receptor-mediated endocytosis: concepts emerging from the LDL receptor system. *Annu Rev Cell Biol* 1: 1–39.
- Goldstein JL, Dana SE, Brown MS. 1974. Esterification of low density lipoprotein cholesterol in human fibroblasts and its absence in homozygous familial hypercholesterolemia. *Proc Natl Acad Sci U S A* 71: 4288–4292.
- Goldstein JL, Hobbs HH, Brown MS. 1995. Familial Hypercholesterolemia. In: Scriver CR, Beaudet AL, Sly WS, and Valle D, editors. *The metabolic and molecular bases of inherited disease.* , 7th Ed.e. New York: McGraw-Hill, p 1981–2030.
- Goldstein JL, Schrott HG, Hazzard WR, Bierman EL, Motulsky AG. 1973. Hyperlipidemia in coronary heart disease. II. Genetic analysis of lipid levels in 176 families and delineation of a new inherited disorder, combined hyperlipidemia. *J Clin Invest* 52: 1544–1568.
- González-Pérez A, López-Bigas N. 2011. Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score, Condel. *Am. J. Hum. Genet.* 88: 440–9.
- Grantham R. 1974. Amino Acid Difference Formula to Help Explain Protein Evolution. *Science* (80-. ). 185: 862–864.



- Group SSC on behalf of the SBR. 1991. Risk of fatal coronary heart disease in familial hypercholesterolaemia. *BMJ* 303: 893–896.
- He G, Gupta S, Yi M, Michaely P, Hobbs HH, Cohen JC. 2002. ARH is a modular adaptor protein that interacts with the LDL receptor, clathrin, and AP-2. *J. Biol. Chem.* 277: 44044–9.
- Hebsgaard SM, Korning PG, Tolstrup N, Engelbrecht J, Rouzé P, Brunak S. 1996. Splice site prediction in *Arabidopsis thaliana* pre-mRNA by combining local and global sequence information. *Nucleic Acids Res.* 24: 3439–52.
- Hobbs HH, Brown MS, Goldstein JL. 1992. Molecular genetics of the LDL receptor gene in familial hypercholesterolemia. *Hum Mutat* 1: 445–466.
- Hobbs HH, Brown MS, Goldstein JL, Russell DW. 1986. Deletion of exon encoding cysteine-rich repeat of low density lipoprotein receptor alters its binding specificity in a subject with familial hypercholesterolemia. *J Biol Chem* 261: 13114–13120.
- Hobbs HH, Russell DW, Brown MS, Goldstein JL. 1990. The LDL receptor locus in familial hypercholesterolemia: mutational analysis of a membrane protein. *Annu Rev Genet* 24: 133–170.
- Holdt LM, Teupser D. 2013. From genotype to phenotype in human atherosclerosis--recent findings. *Curr. Opin. Lipidol.* 24: 410–8.
- Homer VM, Marais AD, Charlton F, Laurie AD, Hurndell N, Scott R, Mangili F, Sullivan DR, Barter PJ, Rye K-A, George PM, Lambert G. 2008. Identification and characterization of two non-secreted PCSK9 mutants associated with familial hypercholesterolemia in cohorts from New Zealand and South Africa. *Atherosclerosis* 196: 659–66.
- Hopkins PN. 2010. Defining the challenges of familial hypercholesterolemia screening: introduction. *J. Clin. Lipidol.* 4: 342–5.
- Hopkins PN, Toth PP, Ballantyne CM, Rader DJ. 2011. Familial hypercholesterolemias: prevalence, genetics, diagnosis and screening recommendations from the National Lipid Association Expert Panel on Familial Hypercholesterolemia. *J. Clin. Lipidol.* 5: S9–17.
- Horton JD, Cohen JC, Hobbs HH. 2007. Molecular biology of PCSK9: its role in LDL metabolism. *Trends Biochem Sci* 32: 71–77.
- Horton JD, Shimomura I. 1999. Sterol regulatory element-binding proteins: activators of cholesterol and fatty acid biosynthesis. *Curr. Opin. Lipidol.* 10: 143–50.
- Hua X, Wu J, Goldstein JL, Brown MS, Hobbs HH. 1995. Structure of the human gene encoding sterol regulatory element binding protein-1 (SREBF1) and localization of SREBF1 and SREBF2 to chromosomes 17p11.2 and 22q13. *Genomics* 25: 667–73.
- Hua X, Yokoyama C, Wu J, Briggs MR, Brown MS, Goldstein JL, Wang X. 1993. SREBP-2, a second basic-helix-loop-helix-leucine zipper protein that stimulates transcription by binding to a sterol regulatory element. *Proc. Natl. Acad. Sci. U. S. A.* 90: 11603–7.
- Hubacek JA, Hyatt T. 2004. ARH missense polymorphisms and plasma cholesterol levels. *Clin. Chem. Lab. Med.* 42: 989–90.

- Huijgen R, Kindt I, Defesche JC, Kastelein JJP. 2012. Cardiovascular risk in relation to functionality of sequence variants in the gene coding for the low-density lipoprotein receptor: a study among 29,365 individuals tested for 64 specific low-density lipoprotein-receptor sequence variants. *Eur. Heart J.* 33: 2325–30.
- Huijgen R, Kindt I, Fouchier SW, Defesche JC, Hutten BA, Kastelein JJP, Vissers MN. 2010. Functionality of sequence variants in the genes coding for the low-density lipoprotein receptor and apolipoprotein B in individuals with inherited hypercholesterolemia. *Hum. Mutat.* 31: 752–60.
- Human Genome Sequencing ConsortiumInternational. 2004. Finishing the euchromatic sequence of the human genome. *Nature* 431: 931–45.
- Ikonen E. 2006. Mechanisms for cellular cholesterol transport: defects and human disease. *Physiol. Rev.* 86: 1237–61.
- Innerarity TL, Weisgraber KH, Arnold KS, Mahley RW, Krauss RM, Vega GL, Grundy SM. 1987. Familial defective apolipoprotein B-100: low density lipoproteins with abnormal receptor binding. *Proc. Natl. Acad. Sci. U. S. A.* 84: 6919–23.
- Jensen HK, Jensen TG, Jensen LG, Hansen PS, Kjeldsen M, Andresen BS, Nielsen V, Meinertz H, Hansen AB, Bolund L. 1994. Characterization of a disease-causing Glu119-Lys mutation in the low-density lipoprotein receptor gene in two Danish families with heterozygous familial hypercholesterolemia. *Hum. Mutat.* 4: 102–13.
- Kathiresan S, Melander O, Guiducci C, Surti A, Burt NP, Rieder MJ, Cooper GM, Roos C, Voight BF, Havulinna AS, Wahlstrand B, Hedner T, et al. 2008. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat. Genet.* 40: 189–97.
- Khachadurian AK. 1964. THE INHERITANCE OF ESSENTIAL FAMILIAL HYPERCHOLESTEROLEMIA. *Am. J. Med.* 37: 402–7.
- Kingston HM. 1989. ABC of clinical genetics. Mendelian inheritance. *BMJ* 298: 375–8.
- Kjolby M, Andersen OM, Breiderhoff T, Fjorback AW, Pedersen KM, Madsen P, Jansen P, Heeren J, Willnow TE, Nykjaer A. 2010. Sort1, encoded by the cardiovascular risk locus 1p13.3, is a regulator of hepatic lipoprotein export. *Cell Metab.* 12: 213–23.
- Knight BL, Gavigan SJ, Soutar AK, Patel DD. 1989. Defective processing and binding of low-density lipoprotein receptors in fibroblasts from a familial hypercholesterolaemic subject. *Eur J Biochem* 179: 693–698.
- Knight BL, Patel DD, Soutar AK. 1987. Regulation of synthesis and cell content of the low-density-lipoprotein receptor protein in cultured fibroblasts from normal and familial hypercholesterolaemic subjects. *Eur J Biochem* 163: 189–196.
- Kotowski IK, Pertsemlidis A, Luke A, Cooper RS, Vega GL, Cohen JC, Hobbs HH. 2006. A spectrum of PCSK9 alleles contributes to plasma levels of low-density lipoprotein cholesterol. *Am. J. Hum. Genet.* 78: 410–22.
- Krisko A, Etchebest C. 2007. Theoretical model of human apolipoprotein B100 tertiary structure. *Proteins* 66: 342–58.

- Ku C-S, Naidoo N, Pawitan Y. 2011. Revisiting Mendelian disorders through exome sequencing. *Hum. Genet.* 129: 351–70.
- Kuczmarski RJ, Ogden CL, Guo SS, Grummer-Strawn LM, Flegal KM, Mei Z, Wei R, Curtin LR, Roche AF, Johnson CL. 2002. 2000 CDC Growth Charts for the United States: methods and development. *Vital Health Stat.* 11. 1–190.
- Kuhlenbäumer G, Hullmann J, Appenzeller S. 2011. Novel genomic techniques open new avenues in the analysis of monogenic disorders. *Hum. Mutat.* 32: 144–51.
- Kusters DM, Huijgen R, Defesche JC, Vissers MN, Kindt I, Hutten BA, Kastelein JJP. 2011. Founder mutations in the Netherlands: geographical distribution of the most prevalent mutations in the low-density lipoprotein receptor and apolipoprotein B genes. *Neth. Heart J.* 19: 175–182.
- Kwiterovich PO, Fredrickson DS, Levy RI. 1974. Familial hypercholesterolemia (one form of familial type II hyperlipoproteinemia). A study of its biochemical, genetic and clinical presentation in childhood. *J. Clin. Invest.* 53: 1237–49.
- Lehrman MA, Russell DW, Goldstein JL, Brown MS. 1987. Alu-Alu recombination deletes splice acceptor sites and produces secreted low density lipoprotein receptor in a subject with familial hypercholesterolemia. *J Biol Chem* 262: 3354–3361.
- Leigh SEA, Foster AH, Whittall RA, Hubbart CS, Humphries SE. 2008. Update and analysis of the University College London low density lipoprotein receptor familial hypercholesterolemia database. *Ann. Hum. Genet.* 72: 485–98.
- Leitersdorf E, Reshef A, Meiner V, Dann EJ, Beigel Y, Roggen FG van, Westhuyzen DR van der, Coetzee GA. 1993. A missense mutation in the low density lipoprotein receptor gene causes familial hypercholesterolemia in Sephardic Jews. *Hum. Genet.* 91: 141–7.
- Lelli N, Ghisellini M, Calandra S, Gaddi A, Ciarrocchi A, Coviello DA, Bertolini S. 1991. Duplication of exons 13, 14 and 15 of the LDL-receptor gene in a patient with heterozygous familial hypercholesterolemia. *Hum. Genet.* 86: 359–62.
- Leren TP, Finborud TH, Manshaus TE, Ose L, Berge KE. 2008. Diagnosis of familial hypercholesterolemia in general practice using clinical diagnostic criteria or genetic testing as part of cascade genetic screening. *Community Genet.* 11: 26–35.
- Leren TP, Solberg K, Rødningen OK, Tonstad S, Ose L. 1994. Two founder mutations in the LDL receptor gene in Norwegian familial hypercholesterolemia subjects. *Atherosclerosis* 111: 175–82.
- Linsel-Nitschke P, Samani NJ, Schunkert H. 2010. Sorting out cholesterol and coronary artery disease. *N. Engl. J. Med.* 363: 2462–3.
- Lodish H, Berk A, Zipursky SL, Baltimore D, Paul M, James D. 2000. *Molecular Cell Biology*, 4th edition. New York: W. H. Freeman.
- Lombardi P, Sijbrands EJ, Giessen K van de, Smelt AH, Kastelein JJ, Frants RR, Havekes LM. 1995. Mutations in the low density lipoprotein receptor gene of familial hypercholesterolemic patients detected by denaturing gradient gel electrophoresis and direct sequencing. *J Lipid Res* 36: 860–867.

- Lombardi P, Sijbrands EJ, Kamerling S, Leuven JA, Havekes LM. 1997. The T705I mutation of the low density lipoprotein receptor gene (FH Paris-9) does not cause familial hypercholesterolemia. *Hum. Genet.* 99: 106–7.
- Ma YH, Bétard C, Roy M, Davignon J, Kessling AM. 1989. Identification of a second “French Canadian” LDL receptor gene deletion and development of a rapid method to detect both deletions. *Clin. Genet.* 36: 219–28.
- Mahley RW, Innerarity TL, Pitas RE, Weisgraber KH, Brown JH, Gross E. 1977. Inhibition of lipoprotein binding to cell surface receptors of fibroblasts following selective modification of arginyl residues in arginine-rich and B apoproteins. *J. Biol. Chem.* 252: 7279–87.
- Mahley RW, Weisgraber KH, Innerarity TL. 1979. Interaction of plasma lipoproteins containing apolipoproteins B and E with heparin and cell surface receptors. *Biochim. Biophys. Acta* 575: 81–91.
- Majewski J, Schwartzentruber J, Lalonde E, Montpetit A, Jabado N. 2011. What can exome sequencing do for you? *J. Med. Genet.* 48: 580–9.
- Mak YT, Pang CP, Tomlinson B, Zhang J, Chan YS, Mak TW, Masarei JR. 1998. Mutations in the low-density lipoprotein receptor gene in Chinese familial hypercholesterolemia patients. *Arterioscler. Thromb. Vasc. Biol.* 18: 1600–5.
- Mancia G, Backer G De, Dominiczak A, Cifkova R, Fagard R, Germano G, Grassi G, Heagerty AM, Kjeldsen SE, Laurent S, Narkiewicz K, Ruilope L, et al. 2007. 2007 Guidelines for the Management of Arterial Hypertension: The Task Force for the Management of Arterial Hypertension of the European Society of Hypertension (ESH) and of the European Society of Cardiology (ESC). *J. Hypertens.* 25: 1105–87.
- Mardis ER. 2008. The impact of next-generation sequencing technology on genetics. *Trends Genet.* 24: 133–41.
- Marduel M, Carrié A, Sassolas A, Devillers M, Carreau V, Filippo M Di, Erlich D, Abifadel M, Marques-Pinheiro A, Munnich A, Junien C, Boileau C, et al. 2010. Molecular spectrum of autosomal dominant hypercholesterolemia in France. *Hum. Mutat.* 31: E1811–24.
- Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen Y-J, Chen Z, Dewell SB, Du L, et al. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437: 376–80.
- Marks D, Thorogood M, Neil HA, Humphries SE. 2003. A review on the diagnosis, natural history, and treatment of familial hypercholesterolaemia. *Atherosclerosis* 168: 1–14.
- Marks D, Wonderling D, Thorogood M, Lambert H, Humphries SE, Neil HA. 2000. Screening for hypercholesterolaemia versus case finding for familial hypercholesterolaemia: a systematic review and cost-effectiveness analysis. *Heal. Technol Assess* 4: 1–123.
- März W, Baumstark MW, Scharnagl H, Ruzicka V, Buxbaum S, Herwig J, Pohl T, Russ A, Schaaf L, Berg A. 1993. Accumulation of “small dense” low density lipoproteins (LDL) in a homozygous patients with familial defective apolipoprotein B-100 results from heterogenous interaction of LDL subfractions with the LDL receptor. *J. Clin. Invest.* 92: 2922–33.

- Maxam AM, Gilbert W. 1977. A new method for sequencing DNA. *Proc. Natl. Acad. Sci. U. S. A.* 74: 560–4.
- Maxwell KN, Breslow JL. 2004. Adenoviral-mediated expression of Pcsk9 in mice results in a low-density lipoprotein receptor knockout phenotype. *Proc. Natl. Acad. Sci. U. S. A.* 101: 7100–5.
- Maxwell KN, Soccio RE, Duncan EM, Sehayek E, Breslow JL. 2003. Novel putative SREBP and LXR target genes identified by microarray analysis in liver of cholesterol-fed mice. *J. Lipid Res.* 44: 2109–19.
- McNutt MC, Lagace TA, Horton JD. 2007. Catalytic activity is not required for secreted PCSK9 to reduce low density lipoprotein receptors in HepG2 cells. *J. Biol. Chem.* 282: 20799–803.
- Medeiros AM, Alves AC, Francisco V, Bourbon M. 2010. Update of the Portuguese Familial Hypercholesterolaemia Study. *Atherosclerosis* 212: 553–558.
- Meng QH, Pajukanta P, Ilmonen M, Schuster H, Schewe CK, Andersson LC, Tikkanen MJ. 1996. Analysis of novel apolipoprotein B mutations using a modified U937 cell line LDL binding assay. *Clin. Chim. Acta.* 256: 27–36.
- Mettlen M, Loerke D, Yarar D, Danuser G, Schmid SL. 2010. Cargo- and adaptor-specific mechanisms regulate clathrin-mediated endocytosis. *J. Cell Biol.* 188: 919–33.
- Metzker ML. 2010. Sequencing technologies - the next generation. *Nat. Rev. Genet.* 11: 31–46.
- Miserez AR, Keller U. 1995. Differences in the phenotypic characteristics of subjects with familial defective apolipoprotein B-100 and familial hypercholesterolemia. *Arterioscler. Thromb. Vasc. Biol.* 15: 1719–29.
- Miserez AR, Muller PY, Barella L, Barella S, Staehelin HB, Leitersdorf E, Kark JD, Friedlander Y. 2002. Sterol-regulatory element-binding protein (SREBP)-2 contributes to polygenic hypercholesterolaemia. *Atherosclerosis* 164: 15–26.
- Mishra SK, Watkins SC, Traub LM. 2002. The autosomal recessive hypercholesterolemia (ARH) protein interfaces directly with the clathrin-coat machinery. *Proc. Natl. Acad. Sci. U. S. A.* 99: 16099–104.
- Motazacker MM, Pirruccello J, Huijgen R, Do R, Gabriel S, Peter J, Kuivenhoven JA, Defesche JC, Kastelein JJP, Hovingh GK, Zelcer N, Kathiresan S, et al. 2012. Advances in genetics show the need for extending screening strategies for autosomal dominant hypercholesterolaemia. *Eur. Heart J.* 33: 1360–6.
- Mozas P, Castillo S, Tejedor D, Reyes G, Alonso R, Franco M, Saenz P, Fuentes F, Almagro F, Mata P, Pocoví M. 2004. Molecular characterization of familial hypercholesterolemia in Spain: identification of 39 novel and 77 recurrent mutations in LDLR. *Hum. Mutat.* 24: 187.
- Muller C. 1938. Xanthomata, hypercholesterolemia, angina pectoris. *Acta Med Scand* 89: 75–84.

- Muller PY, Miserez AR. 2002. Identification of mutations in the gene encoding sterol regulatory element binding protein (SREBP)-2 in hypercholesterolaemic subjects. *J. Med. Genet.* 39: 271–5.
- Musunuru K, Strong A, Frank-Kamenetsky M, Lee NE, Ahfeldt T, Sachs K V, Li X, Li H, Kuperwasser N, Ruda VM, Pirruccello JP, Muchmore B, et al. 2010. From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. *Nature* 466: 714–9.
- Myant NB. 1993. Familial defective apolipoprotein B-100: a review, including some comparisons with familial hypercholesterolaemia. *Atherosclerosis* 104: 1–18.
- Naoumova RP, Neuwirth C, Pottinger B, Whittall R, Humphries SE, Soutar AK. 2004. Genetic diagnosis of familial hypercholesterolaemia: a mutation and a rare non-pathogenic amino acid variant in the same family. *Atherosclerosis* 174: 67–71.
- Naoumova RP, Tosi I, Patel D, Neuwirth C, Horswell SD, Marais AD, Heyningen C van, Soutar AK. 2005. Severe hypercholesterolemia in four British families with the D374Y mutation in the PCSK9 gene: long-term follow-up and treatment response. *Arter. Thromb Vasc Biol* 25: 2654–2660.
- National High Blood Pressure Education Program Working Group on High Blood Pressure in Children and Adolescents. 2004. The fourth report on the diagnosis, evaluation, and treatment of high blood pressure in children and adolescents. *Pediatrics* 114: 555–76.
- Neil HA, Huxley RR, Hawkins MM, Durrington PN, Betteridge DJ, Humphries SE. 2003. Comparison of the risk of fatal coronary heart disease in treated xanthomatous and non-xanthomatous heterozygous familial hypercholesterolaemia: a prospective registry study. *Atherosclerosis* 170: 73–78.
- Ng PC, Henikoff S. 2003. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* 31: 3812–4.
- Ng PC, Kirkness EF. 2010. Whole genome sequencing. *Methods Mol. Biol.* 628: 215–26.
- NICE NCC for PC-. 2008. Identification and management of familial hypercholesterolaemia. NICE clinical guideline 71. . London: National Institute for Health and Clinical Excellence. 45.
- Nigon F, Lesnik P, Rouis M, Chapman MJ. 1991. Discrete subspecies of human low density lipoproteins are heterogeneous in their interaction with the cellular LDL receptor. *J. Lipid Res.* 32: 1741–53.
- Nordestgaard BG, Chapman MJ, Humphries SE, Ginsberg HN, Masana L, Descamps OS, Wiklund O, Hegele RA, Raal FJ, Defesche JC, Wiegman A, Santos RD, et al. 2013. Familial hypercholesterolaemia is underdiagnosed and undertreated in the general population: guidance for clinicians to prevent coronary heart disease: Consensus Statement of the European Atherosclerosis Society. *Eur. Heart J.* 34: 3478–90.
- Norman D, Sun XM, Bourbon M, Knight BL, Naoumova RP, Soutar AK. 1999. Characterization of a novel cellular defect in patients with phenotypic homozygous familial hypercholesterolemia. *J Clin Invest* 104: 619–628.

- Okorodudu DE, Crowley MJ, Sebastian S, Rowell J V, Guyton JR. 2013. Inherited lipemic splenomegaly and the spectrum of apolipoprotein E p.Leu167del mutation phenotypic variation. *J. Clin. Lipidol.* 7: 566–72.
- Online Mendelian Inheritance in Man O (TM). 2000.
- Palacios L, Grandoso L, Cuevas N, Olano-Martín E, Martínez A, Tejedor D, Stef M. 2012. Molecular characterization of familial hypercholesterolemia in Spain. *Atherosclerosis* 221: 137–42.
- Pareek CS, Smoczynski R, Tretyn A. 2011. Sequencing technologies and genome sequencing. *J. Appl. Genet.* 52: 413–35.
- Pereira E, Ferreira R, Hermelin B, Thomas G, Bernard C, Bertrand V, Nassiff H, Mendez del Castillo D, Bereziat G, Benlian P. 1995. Recurrent and novel LDL receptor gene mutations causing heterozygous familial hypercholesterolemia in La Habana. *Hum. Genet.* 96: 319–22.
- Poirier S, Mayer G. 2013. The biology of PCSK9 from the endoplasmic reticulum to lysosomes: new and emerging therapeutics to control low-density lipoprotein cholesterol. *Drug Des. Devel. Ther.* 7: 1135–1148.
- Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. 2010. Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res.* 20: 110–21.
- Pop M, Salzberg SL. 2008. Bioinformatics challenges of new sequencing technology. *Trends Genet.* 24: 142–9.
- Powell LM, Wallis SC, Pease RJ, Edwards YH, Knott TJ, Scott J. 1987. A novel form of tissue-specific RNA processing produces apolipoprotein-B48 in intestine. *Cell* 50: 831–40.
- Pramfalk C, Angelin B, Eriksson M, Parini P. 2007. Cholesterol regulates ACAT2 gene expression and enzyme activity in human hepatoma cells. *Biochem. Biophys. Res. Commun.* 364: 402–9.
- Prassl R, Laggner P. 2009. Molecular structure of low density lipoprotein: current status and future challenges. *Eur. Biophys. J.* 38: 145–58.
- Pullinger CR, Eng C, Salen G, Shefer S, Batta AK, Erickson SK, Verhagen A, Rivera CR, Mulvihill SJ, Malloy MJ, Kane JP. 2002. Human cholesterol 7 $\alpha$ -hydroxylase (CYP7A1) deficiency has a hypercholesterolemic phenotype. *J. Clin. Invest.* 110: 109–17.
- Quail MA, Kozarewa I, Smith F, Scally A, Stephens PJ, Durbin R, Swerdlow H, Turner DJ. 2008. A large genome center's improvements to the Illumina sequencing system. *Nat. Methods* 5: 1005–10.
- Rashid S, Curtis DE, Garuti R, Anderson NN, Bashmakov Y, Ho YK, Hammer RE, Moon Y-A, Horton JD. 2005. Decreased plasma cholesterol and hypersensitivity to statins in mice lacking Pcsk9. *Proc. Natl. Acad. Sci. U. S. A.* 102: 5374–9.
- Raungaard B, Heath F, Brorholt-Petersen JU, Jensen HK, Faergeman O. 1998. Flow cytometry with a monoclonal antibody to the low density lipoprotein receptor compared with gene mutation detection in diagnosis of heterozygous familial hypercholesterolemia. *Clin. Chem.* 44: 966–72.

- Raungaard B, Heath F, Brorholt-Petersen JU, Jensen HK, Faergeman O. 1999. Flow cytometric assessment of LDL receptor activity in peripheral blood mononuclear cells compared to gene mutation detection in diagnosis of heterozygous familial hypercholesterolemia. *Cytometry* 36: 52–9.
- Reese MG, Eeckman FH, Kulp D, Haussler D. 1997. Improved splice site detection in Genie. *J. Comput. Biol.* 4: 311–23.
- Rifai N, Bachorik PS, Albers JJ. 1999. Lipids, Lipoproteins and Apolipoproteins. In: Burtis Ashwood, E. CA, editors. *Text Book of Clinical Chemistry.*, 3rd ed.e. Philadelphia: WB Saunders & Co., p 809–861.
- Rothblat G. 1969. Lipid metabolism in tissue culture cells. *Adv Lipid Res.* 7: 135–162.
- Rubinsztein DC, Jialal I, Leitersdorf E, Coetzee GA, Westhuyzen DR van der. 1993. Identification of two new LDL-receptor mutations causing homozygous familial hypercholesterolemia in a South African of Indian origin. *Biochim. Biophys. Acta* 1182: 75–82.
- Rudenko G, Henry L, Henderson K, Ichtchenko K, Brown MS, Goldstein JL, Deisenhofer J. 2002. Structure of the LDL receptor extracellular domain at endosomal pH. *Science* 298: 2353–8.
- Sandhofer A, Tatarczyk T, Kirchmair R, Iglseder B, Paulweber B, Patsch JR, Schratzberger P. 2009. Are plasma VEGF and its soluble receptor sFlt-1 atherogenic risk factors? Cross-sectional data from the SAPHIR study. *Atherosclerosis* 206: 265–9.
- Sandhu MS, Waterworth DM, Debenham SL, Wheeler E, Papadakis K, Zhao JH, Song K, Yuan X, Johnson T, Ashford S, Inouye M, Luben R, et al. 2008. LDL-cholesterol concentrations: a genome-wide association study. *Lancet* 371: 483–91.
- Sanger F, Nicklen S, Coulson AR. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* 74: 5463–7.
- Sato R, Takano T. 1995. Regulation of intracellular cholesterol metabolism. *Cell Struct. Funct.* 20: 421–7.
- Schneider WJ, Beisiegel U, Goldstein JL, Brown MS. 1982. Purification of the low density lipoprotein receptor, an acidic glycoprotein of 164,000 molecular weight. *J. Biol. Chem.* 257: 2664–73.
- Schwarz JM, Rödelberger C, Schuelke M, Seelow D. 2010. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat. Methods* 7: 575–6.
- Segrest JP, Jones MK, Loof H De, Dashti N. 2001. Structure of apolipoprotein B-100 in low density lipoproteins. *J. Lipid Res.* 42: 1346–67.
- Seidah NG, Benjannet S, Wickham L, Marcinkiewicz J, Jasmin SB, Stifani S, Basak A, Prat A, Chretien M. 2003. The secretory proprotein convertase neural apoptosis-regulated convertase 1 (NARC-1): liver regeneration and neuronal differentiation. *Proc. Natl. Acad. Sci. U. S. A.* 100: 928–33.
- Shendure J, Ji H. 2008. Next-generation DNA sequencing. *Nat. Biotechnol.* 26: 1135–45.



- Silva S, Alves AC, Patel D, Malho R, Soutar AK, Bourbon M. 2012. In vitro functional characterization of missense mutations in the LDLR gene. *Atherosclerosis*.
- Sjouke B, Kusters DM, Kindt I, Besseling J, Defesche JC, Sijbrands EJG, Roeters van Lennep JE, Stalenhoef AFH, Wiegman A, Graaf J de, Fouchier SW, Kastelein JJP, et al. 2014. Homozygous autosomal dominant hypercholesterolaemia in the Netherlands: prevalence, genotype-phenotype relationship, and clinical outcome. *Eur. Heart J*.
- Soria LF, Ludwing EH, Clarke HRG, Vega GL, Grundy SM, McCarthy BJ. 1989. Association between a specific apolipoprotein B mutation and familial defective apolipoprotein B-100. *Proc Natl Acad Sci U S A* 86: 587–591.
- Soutar AK. 2010. Rare genetic causes of autosomal dominant or recessive hypercholesterolaemia. *IUBMB Life* 62: 125–131.
- Soutar AK, Knight BL, Patel DD. 1989. Identification of a point mutation in growth factor repeat C of the low density lipoprotein-receptor gene in a patient with homozygous familial hypercholesterolemia that affects ligand binding and intracellular movement of receptors. *Proc. Natl. Acad. Sci. U. S. A.* 86: 4166–70.
- Soutar AK, Naoumova RP. 2004. Autosomal recessive hypercholesterolemia. *Semin Vasc Med* 4: 241–248.
- Soutar AK, Naoumova RP. 2007. Mechanisms of disease: genetic causes of familial hypercholesterolemia. *Nat Clin Pr. Cardiovasc Med* 4: 214–225.
- Soutar AK, Naoumova RP, Traub LM. 2003. Genetics, clinical phenotype, and molecular cell biology of autosomal recessive hypercholesterolemia. *Arter. Thromb Vasc Biol* 23: 1963–1970.
- Srivastava A, Srivastava A, Srivastava K, Choudhuri G, Mittal B. 2010. Role of ABCG8 D19H (rs11887534) variant in gallstone susceptibility in northern India. *J. Gastroenterol. Hepatol.* 25: 1758–62.
- Südhof TC, Goldstein JL, Brown MS, Russell DW. 1985. The LDL receptor gene: a mosaic of exons shared with different proteins. *Science* 228: 815–22.
- Sun XM, Patel DD, Knight BL, Soutar AK. 1998. Influence of genotype at the low density lipoprotein (LDL) receptor gene locus on the clinical phenotype and response to lipid-lowering drug therapy in heterozygous familial hypercholesterolaemia. The Familial Hypercholesterolaemia Regression Study Group. *Atherosclerosis* 136: 175–85.
- Talmud PJ, Shah S, Whittall R, Futema M, Howard P, Cooper JA, Harrison SC, Li K, Drenos F, Karpe F, Neil HAW, Descamps OS, et al. 2013. Use of low-density lipoprotein cholesterol gene score to distinguish patients with polygenic and monogenic familial hypercholesterolaemia: a case-control study. *Lancet* 381: 1293–301.
- Taschner PEM, Dunnen JT den. 2011. Describing structural changes by extending HGVS sequence variation nomenclature. *Hum. Mutat.* 32: 507–11.
- Taylor A, Martin B, Wang D, Patel K, Humphries SE, Norbury G. 2009. Multiplex ligation-dependent probe amplification analysis to screen for deletions and duplications of the LDLR gene in patients with familial hypercholesterolaemia. *Clin. Genet.* 76: 69–75.

- Taylor A, Tabrah S, Wang D, Sozen M, Duxbury N, Whittall R, Humphries SE, Norbury G. 2007. Multiplex ARMS analysis to detect 13 common mutations in familial hypercholesterolaemia. *Clin. Genet.* 71: 561–8.
- Teer JK, Mullikin JC. 2010. Exome sequencing: the sweet spot before whole genomes. *Hum. Mol. Genet.* 19: R145–51.
- Thomas ERA, Atanur SS, Norsworthy PJ, Encheva V, Snijders AP, Game L, Vandrovcova J, Siddiq A, Seed M, Soutar AK, Aitman TJ. 2013. Identification and biochemical analysis of a novel APOB mutation that causes autosomal dominant hypercholesterolemia. *Mol. Genet. Genomic Med.* 1: 155–161.
- Tosi I, Toledo-Leiva P, Neuwirth C, Naoumova RP, Soutar AK. 2007. Genetic defects causing familial hypercholesterolaemia: identification of deletions and duplications in the LDL-receptor gene and summary of all mutations found in patients attending the Hammersmith Hospital Lipid Clinic. *Atherosclerosis* 194: 102–111.
- Tsiatis AC, Norris-Kirby A, Rich RG, Hafez MJ, Gocke CD, Eshleman JR, Murphy KM. 2010. Comparison of Sanger sequencing, pyrosequencing, and melting curve analysis for the detection of KRAS mutations: diagnostic and clinical implications. *J. Mol. Diagn.* 12: 425–32.
- Tveten K, Strøm TB, Cameron J, Berge KE, Leren TP. 2012. Mutations in the SORT1 gene are unlikely to cause autosomal dominant hypercholesterolemia. *Atherosclerosis* 225: 370–5.
- Usifo E, Leigh SE, Whittall RA, Lench N, Taylor A, Yeats C, Orengo CA, Martin AC, Celli J, Humphries SE. 2012. Low-density lipoprotein receptor gene familial hypercholesterolemia variant database: update and pathological assessment. *Ann Hum Genet* 76: 387–401.
- Usui R, Shibuya M, Ishibashi S, Maru Y. 2007. Ligand-independent activation of vascular endothelial growth factor receptor 1 by low-density lipoprotein. *EMBO Rep.* 8: 1155–61.
- Vega GL, Grundy SM. 1986. In vivo evidence for reduced binding of low density lipoproteins to receptors as a cause of primary moderate hypercholesterolemia. *J. Clin. Invest.* 78: 1410–4.
- Voelkerding K V, Dames SA, Durtschi JD. 2009. Next-generation sequencing: from basic research to diagnostics. *Clin. Chem.* 55: 641–58.
- Vrablík M, Ceska R, Horínek A. 2001. Major apolipoprotein B-100 mutations in lipoprotein metabolism and atherosclerosis. *Physiol. Res.* 50: 337–43.
- Wang J, Ban MR, Hegele RA. 2005. Multiplex ligation-dependent probe amplification of LDLR enhances molecular diagnosis of familial hypercholesterolemia. *J. Lipid Res.* 46: 366–72.
- Weisgraber KH, Innerarity TL, Mahley RW. 1978. Role of lysine residues of plasma lipoproteins in high affinity binding to cell surface receptors on human fibroblasts. *J. Biol. Chem.* 253: 9053–62.
- Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, He W, Chen Y-J, Makhijani V, Roth GT, Gomes X, Tartaro K, et al. 2008. The complete genome of an individual by massively parallel DNA sequencing. *Nature* 452: 872–6.

- Whitfield AJ, Barrett PHR, Bockxmeer FM van, Burnett JR. 2004. Lipid disorders and mutations in the APOB gene. *Clin. Chem.* 50: 1725–32.
- WHO. 1998.
- Willer CJ, Sanna S, Jackson AU, Scuteri A, Bonnycastle LL, Clarke R, Heath SC, Timpson NJ, Najjar SS, Stringham HM, Strait J, Duren WL, et al. 2008. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat. Genet.* 40: 161–9.
- Yamakawa K, Takada K, Yanagi H, Tsuchiya S, Kawai K, Nakagawa S, Kajiyama G, Hamaguchi H. 1989. Three novel partial deletions of the low-density lipoprotein (LDL) receptor gene in familial hypercholesterolemia. *Hum. Genet.* 82: 317–21.
- Yamamoto T, Davis CG, Brown MS, Schneider WJ, Casey ML, Goldstein JL, Russell DW. 1984. The human LDL receptor: a cysteine-rich protein with multiple Alu sequences in its mRNA. *Cell* 39: 27–38.
- Yang MQ, Athey BD, Arabnia HR, Sung AH, Liu Q, Yang JY, Mao J, Deng Y. 2009. High-throughput next-generation sequencing technologies foster new cutting-edge computing techniques in bioinformatics. *BMC Genomics* 10 Suppl 1: I1.
- Yoon JH, Kuver R, Choi HS. 2010. ABCG8 D19H polymorphism: a basis for the genetic prediction of cholesterol gallstone disease. *J. Gastroenterol. Hepatol.* 25: 1713–4.
- Yoshimura H, Iwasaki S, Nishio S-Y, Kumakawa K, Tono T, Kobayashi Y, Sato H, Nagai K, Ishikawa K, Ikezono T, Naito Y, Fukushima K, et al. 2014. Massively parallel DNA sequencing facilitates diagnosis of patients with usher syndrome type 1. *PLoS One* 9: e90688.
- Yu L, Gupta S, Xu F, Liverman ADB, Moschetta A, Mangelsdorf DJ, Repa JJ, Hobbs HH, Cohen JC. 2005. Expression of ABCG5 and ABCG8 is required for regulation of biliary cholesterol secretion. *J. Biol. Chem.* 280: 8742–7.
- Yue P, Aversa M, Lin X, Schonfeld G. 2006. The c.43\_44insCTG variation in PCSK9 is associated with low plasma LDL-cholesterol in a Caucasian population. *Hum. Mutat.* 27: 460–6.
- Zhang J, Chiodini R, Badr A, Zhang G. 2011. The impact of next-generation sequencing on genomics. *J. Genet. Genomics* 38: 95–109.
- Zuliani G, Arca M, Signore A, Bader G, Fazio S, Chianelli M, Bellosta S, Campagna F, Montali A, Maioli M, Pacifico A, Ricci G, et al. 1999. Characterization of a new form of inherited hypercholesterolemia: familial recessive hypercholesterolemia. *Arterioscler. Thromb. Vasc. Biol.* 19: 802–9.



**ANEXOS**

---



### Anexo A.I – Sequência dos *primers* utilizados, tamanho do fragmento e temperatura de hibridação para os genes *LDLR* e *APOB*

Exão	5' -3' sequence	tamanho fragmento (bp)	Temperatura de hibridação	Primers
Promotor & Exão 1	F: GGG TTA AAA AGC CGA TGT CA R: GCC ATT ACC CCA CAA GTC TCC	480	59°	SPr+1F SPr+1R
exão 2	F: TCC CAT ACC CCA GAG AGT CCA TA R: CAG CCG CCA TCA TCA AAA AG	587	58°	R2F R2R
exão 3	F: GGT TTC ACT ATA TTG GCC AGG R: CTC CCC AGG ACT CAG ATA GG	327	59°	LDL 3F1 MB260
exão 4	F: GAG GAA ACT GAG GCA CCG AG R: TTG GCA TGT TGT TGG AAA TCC	677	57°	EX4F EX4R
exão 5	F: GCA AAA GGC CCT GCT TCT TT R: GAG GCT CTG AGA AGT CAA GT	342	58°	EX5F NEW EX5R NEW
exão 6	F: TGAATGAGTGCCAAGCAAAC R: TTCCCAAAACCCTACAGCAC	277	59°	MB328 MB329
exão 7	F: GCGAAGGGATGGGTAGGG R: GCATGAGGGGTTTGGTTG	248	58°	MB316 MB317
exão 8	F: ATC TCC CGA GAG GCT GGG CTG TCT R: CCC GGT CAG GGG ATA TGA GTC TGT	361	59°	MB30 MB31
exão 9	F: AAG GGG ATG GGG AGG CAC TCT TG R: CCT CAT CTC ACC TGC GGG CCA	397	59°	EX9+10F MB277
exão 10	F: CCT TGG CCC GCA GGT GAG A R: GTG CTG GGA TTA CAG GTG CTT TGA	403	62°	MB34 MB35
exão 11	F: GCC ACA TTT GGA GTT TGG GGT TC R: AGC AGC TTG GGC TTG TCC CAG A	355	60°	EX11F EX11R
exão 12	F: GGT GCT TTT CTG CTA GGT CC R: TTT TCT GCG TTC ATC TTG GCT	347	59°	EX12F EX12R
exão 13	F: CTA GTT GTG GAG AGA GGG TGG C R: GCG GAG TCA GGG CAG GAA CGA G	275	60°	EX13F EX13R
exão 14	F: GAA ACC TCC TTG TGG AAA CTC T R: GAA AAG TAT GGT TAT CCC GAC T	388	58°	EX14F EX14R
exão 15	F: CCA AGG TCA TTT GAG ACT TTC GT R: GAG AGA AGG TCA GCA AGG GAG TG	388	60°	EX15F EX15R
exão 16	F: GTC CTC TGC CTG CTC CAT TTC TT R: ATC CTC CAT CTG ACC CCT TAG C	350	60°	EX16F EX16R
exão 17	F: GAG CTG GGT CTC TGG TCT CG R: GCG CAC AGA AGC ATT CAC CT	500	60°	R17F R17R
exão 18	F: GAG CGG TGG GAA GTG ACT GAA T R: TGG TGC CAT CTG CTG TTG TGT G	580	59°	EX18F EX18R
ApoB (exão 26)	F: GAGCAGTTGACCACAAGCTTAGCTTGGA R: GGGTGGCTTTGCTTGTATGTTCTCCGT	343	59°	P61 P62
ApoB (exão 29)	F: CCAAGATGAGATCAACACAATC R: AACTTGACTTGAGAGTTGGG	334	59	MB63 MB64

Anexo A.2 – Sequência dos *primers* utilizados, tamanho do fragmento e temperatura de hibridação para o gene *PCSK9*.

Exão	5' -3' Sequence	Tamanho fragmento (bp)	Temperatura de hibridação	Primer
Promotor & Exão 1	F: GGA GGG CGA GGC CGA AAC CTG AT R: CTC GGC GAC CTG CAC TCC ACT TCC	760	63°	MB1 MB3
Exão 2	F: GCT GGG TTT CTT CCA TGT CAT CAT R: CCG ATA AGT GCT CAA TAC ATA CTT GCT	340	54.5°	MB261 MB262
Exão 3	F: CTCTATGCCAGACCGTGTTG R: GTGCTGAGTCCCAAAGCC	400	55°	890 891
Exão 4 & 5	F: AGGTGACTTTATGCTCATTCCT R: TAGGAGACATTAGCTCTCCCTGG	740	60°	MB263 MB264
Exão 6	F: TCGCAGCAGCATTTCAC R: TCCAAAGCCAGAAGGGTTC	400	57°	892 893
Exão 7	F: GGCCTGAGTCTGCCTCTGC R: ACCCTGACTGCCAAAGGGGC	330	65°	MB98 MB99
Exão 8	F: CACTGGCAGGAGTCCCCTGC R: AGGTCACACAGACCTCCCAAGC	320	64°	MB100 MB101
Exão 9	F: GTAAGGAGGATGATGCCACC R: TTACAGAAGAGCTGGAGTCTGG	450	63°	MB199 MB200
Exão 10	F: AGCTCCTGTGCCCCAGAAG R: GAGTATGGAAGTCAAGTCAGG	400	57°	896 897
Exão 11	F: GGCTCAGAGAGGTGAATGG R: GCATCTACCTGGCAAACCG	340	62°	MB6 MB7
Exão 12	F: GTGGGAGATACACGGTTGTGTCC R: TGGGGAGGAGGCACCCAGAGT	400	58°	MB330 MB331

Anexo A.3 – Sequência dos *primers* utilizados, tamanho do fragmento e temperatura de hibridação para RT-PCR

Exão	5' -3' Sequence	Tamanho fragmento (bp)	Temperatura de hibridação	Primer
Exão 2	CAGCGCTGAGTGCCAGGATG	312	62°	MB8
Exão 4	CTGGCGGGACCACAGGTGAG			MB9
Exão 4	CCGCAGCGCTGTAGGGGTCT			MB10
Exão 6	TGACCAGTCCCGGCAGTCTC	319	62°	MB11
Exão 6	ACTGCCGGGACTGGTCAGAT			MB12
Exão 9	GCCGGTTGGTGAAGAAGAGG			MB13
Exão 8	GCGTGAACCTGGAGGGTGGC	325	62°	MB14
Exão 10	GGGGCCTGGATGTCTCTGCT			MB15
Exão 11	GGAACCTCCCGCCAAGATCAA			MB16
Exão 13	TTGGCTGGGTGAGGTTGTGG	376	62°	MB17
Exão 13	TTCAAGTGCCAACCGCCTCAC			MB18
Exão 15	TCGGGAACAGGTTCGGGTGGT			MB19
Exão 15	GGGGCCACCCCTGGGCTCAC	380	62°	MB20
Exão 18	AAGGCCGGCGAGGTCTCAGG			MB21



Anexo A.4 - Referências das sequências nucleóticas e proteicas dos genes e proteínas analisados neste trabalho.

Gene	Protein	UNIPROT	Transcript code
DHCR7	7-dehydrocholesterol reductase	Q9UBM7	NM_001163817
OSBPL1A	Oxysterol-binding protein 1	Q8K2D2	NM_080597
ABCB10	ATP-binding cassette sub-family B member 10, mitochondrial	Q9NRK6	NM_012089
GGPS1	Geranylgeranyl pyrophosphate synthase	O95749	NM_001037277
LRP5	Low-density lipoprotein receptor-related protein 5	O75197	NM_002335
CPT1A	Carnitine O-palmitoyltransferase 1, liver isoform	P50416	NM_001031847
UCP2	Mitochondrial uncoupling protein 2	P55851	NM_003355
UCP3	Mitochondrial uncoupling protein 3	P55916	NM_003356.3
ATI/ ACAT1	Sterol O-acyltransferase 1	P35610	NM_003101.5
APOA1	Apolipoprotein A-I	P02647	NM_000039.1
APOC3	Apolipoprotein C-III	P02656	NM_000040.1
APOA4	<b>Apolipoprotein A-IV</b>	P06727	NM_000482.3
APOA5	Apolipoprotein A-V	Q6Q788	NM_052968.4
CYP11A	<b>Cholesterol side-chain cleavage enzyme, mitochondrial</b>	P05108	NM_001163086.1
LRP3	Low-density lipoprotein receptor-related protein 3	O75074	NM_002333.3
LIPE	hormone-sensitive lipase	Q05469	NM_005357.3
APOC1	Apolipoprotein C-I	P02654	NM_001645.3
APOC2	<b>Apolipoprotein C-II</b>	P02655	NM_000483.4
APOC4	Apolipoprotein C-IV	P55056	NM_001646.2
APOE	<b>Apolipoprotein E</b>	P02649	NM_000041.2
PSRC1	Proline/serine-rich coiled-coil protein 1	Q6PGN9	NM_032636.7
CELSR2	Cadherin EGF LAG seven-pass G-type receptor 2	Q9HCU4	NM_001408
SORT1	Sortilin	Q99523	NM_001205228
NCAN	Neurocan core protein	O14594	NM_004386.2
CILP2	Cartilage intermediate layer protein 2	Q8IUL8	NM_153221
B3GALT4	Beta-1,3-galactosyltransferase 4	O96024	NM_003782.3
B4GALT4	Beta-1,4-galactosyltransferase 4	O60513	NM_212543.1
ABCG8	ATP-binding cassette sub-family G member 8	Q9H221	NM_022437
PBX4	Pre-B-cell leukemia transcription factor 4	Q9BYU1	NM_001024954.1
HMGCR	3-hydroxy-3-methylglutaryl-coenzyme A reductase	P54646	NM_001130996
TIMD4	T-cell immunoglobulin and mucin domain-containing protein 4	Q96H15	NM_001146726
HNF1A	Hepatocyte nuclear factor 1-alpha	P20823	NM_000545
BCL3	B-cell lymphoma 3 protein	P20749	NM_005178.4
TOMM40	Mitochondrial import receptor subunit TOM40 homolog	O96008	NM_006114.2
ABCG1	ATP-binding cassette sub-family G member 1	P45844	NM_207629.1
ABCA1	ATP-binding cassette sub-family A (ABC1), member 1	O95477	NM_005502
LSS	Lanosterol synthase	P48449	NM_001145437
GCKR	Glucokinase regulatory protein	Q14397	NM_001486
MTTP	Microsomal triglyceride transfer protein large subunit	Q6P5T3	NM_000253
SMAD3	Mothers against decapentaplegic homolog 3	P84022	NM_001145102
LRP2	low density lipoprotein receptor-related protein 2	P98164	NM_001174081
LRPAP1	Low Density Lipoprotein Receptor-Related Protein Associated Protein 1	P30533	NM_002337
PDIA3	protein disulfide isomerase family A, member 3	P30101	NM_005313
DAB2	disabled homolog 2, mitogen-responsive phosphoprotein	Q5VWQ8	NM_001244871
FLT1	Vascular endothelial growth factor receptor 1 precursor	P17948	NM_002019